

Lecture Notes in Networks and Systems 321

Milan Tuba  
Shyam Akashe  
Amit Joshi *Editors*

# ICT Systems and Sustainability

Proceedings of ICT4SD 2021, Volume 1

 Springer

# Lecture Notes in Networks and Systems

Volume 321

## Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences,  
Warsaw, Poland

## Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA,  
School of Electrical and Computer Engineering—FEEC, University of Campinas—  
UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering,  
Bogazici University, Istanbul, Turkey

Derong Liu, Department of Electrical and Computer Engineering, University  
of Illinois at Chicago, Chicago, USA

Institute of Automation, Chinese Academy of Sciences, Beijing, China

Witold Pedrycz, Department of Electrical and Computer Engineering, University of  
Alberta, Alberta, Canada

Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Marios M. Polycarpou, Department of Electrical and Computer Engineering,  
KIOS Research Center for Intelligent Systems and Networks, University of Cyprus,  
Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary

Jun Wang, Department of Computer Science, City University of Hong Kong,  
Kowloon, Hong Kong

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at <https://link.springer.com/bookseries/15179>

Milan Tuba · Shyam Akashe · Amit Joshi  
Editors

# ICT Systems and Sustainability

Proceedings of ICT4SD 2021, Volume 1



*Editors*

Milan Tuba  
Singidunum University  
Belgrade, Serbia

Shyam Akashe   
ITM University  
Gwalior, Madhya Pradesh, India

Amit Joshi  
Global Knowledge Research Foundation  
Ahmedabad, Gujarat, India

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-16-5986-7

ISBN 978-981-16-5987-4 (eBook)

<https://doi.org/10.1007/978-981-16-5987-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.  
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Preface

The Sixth International Conference on ICT for Sustainable Development (ICT4SD 2021) targets theory, development, applications, experiences, and evaluation of interaction sciences with fellow students, researchers, and practitioners.

The conference may concern any topics within the conference scope. Workshops may be related to any topics within the conference scope. The conference is devoted to increase the understanding role of technology issues and how engineering has day by day evolved to prepare human-friendly technology. The conference will provide a platform for bringing forth significant research and literature across the field of ICT for sustainable development and provide an overview of the technologies awaiting unveiling. This interaction will be the focal point for leading experts to share their insights, provide guidance, and address participant's questions and concerns.

The conference was to be held during August 5–6, 2021, at Hotel Vivanta, Panaji, Goa, India, but due to the pandemic, this year it was organized through digital mode and few of the papers were presented physically at the venue on August 6, 2021. The conference was organized by the Global Knowledge Research Foundation, Supporting Partner InterYIT, International Federation for Information Processing, State Chamber Partner Goa Chamber of Commerce & Industry, and National Chamber Partner Knowledge Chamber of Commerce & Industry.

Research submissions in various advanced technology areas were received after a rigorous peer review process with the help of program committee members and 193 external reviewers for 850+ papers from 16 different countries including Algeria, the USA, United Arab Emirates, Serbia, Qatar, Mauritius, Egypt, Saudi Arabia, Ethiopia, Oman out of which 180 were accepted with an acceptance ratio of 0.17.

Technology is the driving force of progress in this era of globalization. Information and communication technology (ICT) has become a functional requirement for the socioeconomic growth and sustained development of any country. The influence of information communications technology (ICT) in shaping the process of globalization, particularly in productivity, commercial, and financial spheres, is widely recognized. The ICT sector is undergoing a revolution that has momentous implications for the current and future social and economic situation of all the countries in the world. ICT plays a pivotal role in empowering people for self-efficacy and how

it can facilitate this mission to reach out to grassroots level. Finally, it is concluded that ICT is a significant contributor to the success of the ongoing initiative of Startup India.

In order to recognize and reward the extraordinary performance and achievements by ICT and allied sectors and promote universities, researchers and students through their research work adapting new scientific technologies and innovations. The two-day conference had presentations from the researchers, scientists, academia, and students on the research work carried out by them in different sectors.

The summit will be addressed by eminent dignitaries including Shri Nitin Kunkolienker, President, MAIT; Shri. Prajyot Mainkar, Chairman, IT Committee of Goa Chamber of Commerce and Industry; Mike Hinchey, President, IFIP, and Chair, IEEE, UK and Ireland; Milan Tuba, Vice Rector for International Relations, Singidunum University, Serbia; Prof. Jagdish Bansal, India; Mr. Aninda Bose, Springer; and Dr. Amit Joshi, Director, GR Foundation. The overall conference had one inaugural session, one keynote session, and 18 technical sessions during these two days.

Belgrade, Serbia  
Gwalior, India  
Ahmedabad, India

Milan Tuba  
Shyam Akashe  
Amit Joshi

# Contents

<b>Literature Survey for MIMO Systems and Interference Cancellation in MIMO</b> .....	1
Varshini Rajesh, Shivani Krishna Kompella, and A. R. Abdul Rajak	
<b>Design and Study the Performance of Micro-Strip Patch Antennae for 5G Mobile Communication</b> .....	11
Shivani Krishna Kompella and A. R. Abdul Rajak	
<b>Sentiment Analysis of COVID-19 Vaccine Rollout in India</b> .....	21
Sushila Paliwal, Suraiya Parveen, M. Afshar Alam, and Jawed Ahmed	
<b>Detection of Swine Flu System Using IOT and Machine Learning</b> .....	35
Devyani Bhamare and V. V. Mandhare	
<b>Text Mining and Machine Learning Algorithms to Identifying Diseases and Providing Repair Action Using ICD-10 Codes</b> .....	45
Ashish P. Ramdasi and S. Sathyalakshmi	
<b>Communication Skills: An Integral Instrument for Data Analysis in the Field of Research</b> .....	55
Shweta Arora, Kavita Ajay Joshi, Preet Nain Kaur Narula, Deepa Nainwal, Narendra Bisht, Girish Nainwal, and Ayush Agarwal	
<b>Model for Intrusion Detection in Cyber-Physical System to Address Network Simulation</b> .....	65
R. Aishwarya, M. Ajitha, and A. Sheryl Oliver	
<b>Comparative Study to Perform Tweets/NEWS Classification Using Big Data Analytics for Predicting Stock Market Price Movement</b> .....	75
Dhara N. Darji, Satyen M. Parikh, and Hiral R. Patel	
<b>Design and Improvement in the Performance of Multi-band House-Shaped Microstrip Patch Antenna with SRR-Based Metamaterial for 5G Applications</b> .....	85
John Colaco and Rajesh B. Lohani	

<b>Family of Friends—A Hostel Utility System</b> .....	95
Mahima Badjate, Saniya Ponkshe, Sneha Rohida, Toral Patel, and Prajakta Deshpande	
<b>Vulnerability Due to Training Order in Split Learning</b> .....	103
Harshit Madaan, Manish Gawali Viraj Kulkarni, and Aniruddha Pant	
<b>Efficient Attendance Management System Based on Face Recognition</b> .....	113
Ujwalla Gawande, Pratyush Joshi, Sumedh Ghatwai, Shreyas Nemade, Soham Balkothe, and Nishant Shrikhande	
<b>Automated Medical Symptoms Summarization and Patient Profiling System</b> .....	123
Sakshi Gehani and Siba Panda	
<b>Smart Shoe Provides Vision to Visionless Person</b> .....	131
Roshani K. Dharme, Jaya R. Surywanshi, Harsha C. Kunwar, and Yogesh H. Palve	
<b>Relevance of Disruptive Technologies Led Knowledge Management System and Practices for MSME</b> .....	139
Ravi Shankar Jha, Priti Ranjan Sahoo, and Smrutirekha	
<b>Empirical Analysis of Static and Dynamic Stopword Generation Approaches</b> .....	149
SaziyaBegum Saiyed and Priti Sajja	
<b>Rethinking Access Control Mechanism for Ubiquitous Computing</b> .....	157
Nalini Mhetre, A. V. Deshpande, and Parikshit N. Mahalle	
<b>Review of Smart Building Management System</b> .....	167
Mona Tatari, Parul Agarwal, M. Afshar Alam, and Jawed Ahmed	
<b>Optimal Packet Routing with Multiple Demands and Link Lines</b> .....	177
Monu and L. N. Das	
<b>Mapping and Change Detection of Mangroves Using Remote Sensing and Google Earth Engine: A Case Study</b> .....	187
Geetanjali S. Mahamunkar, Arvind W. Kiwelekar, and Laxman D. Netak	
<b>Source Code Obfuscation: Novel Technique and Implementation</b> .....	197
Krishna Yadav, Rutuja Kamble, Sumit Kale, Yash Pazare, and Pallavi Ahire	
<b>Design and Implementation of Laptop Tracking System Based on Cloud Computing and IoT</b> .....	205
Soumyadip Chatterjee	

<b>Design and Implementation of Women Security System Based on GSM and GPS Technology</b> .....	217
Soumyadip Chatterjee and Surajit Mandal	
<b>A Scalable Framework for High-Performance Computing with Cloud</b> .....	225
Manish Kumar Abhishek and D. Rajeswara Rao	
<b>Mango (Mangifera Indica L.) Classification Using Shape, Texture and Convolutional Neural Network Features</b> .....	237
Sapan Naik, Yash Rana, and Viraj Thakkar	
<b>Modeling Human Tendencies for Password Guessing</b> .....	247
Rohit Mutalik, Dhairya Chheda, Zeeshan Shaikh, and Dhanashree Toradmalle	
<b>Leveraging Big Data Analytics for Predictive Policing in India</b> .....	259
Kashika Wanchoo and Jaimin Desai	
<b>An Assured Encryption Strategy for Private Data Protection in Hybrid Cloud</b> .....	271
D. I. George Amalarethinam and J. Madhupriya	
<b>Reducing Labelled Data Requirement for Pneumonia Segmentation Using Image Augmentations</b> .....	281
Jitesh Seth, Rohit Lokwani, Viraj Kulkarni, Aniruddha Pant, and Amit Kharat	
<b>Feature Extraction for Automatic Grading of Students' Presentations</b> .....	293
Jyoti G. Borade, Arvind W. Kiwelekar, and Laxman D. Netak	
<b>Monitoring and Assessment of Water Bodies in Madurai Using Geo-Spatial Data</b> .....	303
R. Shanthi Priya, R. Sarunisha, S. Aditya, M. Haripriya, and M. Bhuvaneshwari	
<b>Applications of Open Source Intelligence in Crisis Analysis—A COVID-19 Case Study</b> .....	313
A. K. Buvasri, R. Meenakshi, and S. Karthika	
<b>Language Agnostic Generation of Syntactically Incorrect Programs</b> ....	327
Ramayani Ghosh, Rasya Ramesh, V. S. Meghana, and N. S. Kumar	
<b>Distributed FBG Temperature Sensor for Coal Mine Fire Detection</b> ....	337
Sagupha Parween and Aruna Tripathy	
<b>Image Captioning Using Ensemble Model</b> .....	345
Pranav Bathija, Harsh Chawla, Ashish Bhat, and Arti Deshpande	

<b>Graph-Based Transductive Label Propagation Approach for Multi-label Classification</b> .....	357
S. C. Dharmadhikari, S. H. Chandak, and R. R. Chhajed	
<b>Detection and Classification of Diseases and Maturity of Dragon Fruits</b> .....	365
Vrushali Kulkarni, Pranali Kosamkar, Chanavi Singh, Purva Ingle, and Vedant Modi	
<b>Implementation of DevOps Culture in Customizing Serverless Computing and Function as a Service (FaaS) for IoT Mashup</b> .....	375
M. Ganeshan and A. Rajesh	
<b>A Cloud Interfaced Social Ski Driver Optimized Routing for Future Electric Motor Vehicles</b> .....	385
Shivanand C. Hiremath and Jayashree D. Mallapur	
<b>Cloud-Enabled Education-as-a-Service (EaaS)—A Review</b> .....	397
Shubhpreet Kaur, Tarandeep Kaur, and Ashok Sharma	
<b>Blockchain and Cloud Technology: Leading the ICT Innovations</b> .....	405
Shilpa and Tarandeep Kaur	
<b>Automated Gate System Using Number Plate Recognition (NPR)</b> .....	413
Shreya Anekar, Samrudhi Yeginwar, and Harsha Sonune	
<b>Securing the COVID Patients' Medical Records Using Encrypted Image Steganography</b> .....	421
Asha Durafe and Vinod Patidar	
<b>Text Summarization of News Articles</b> .....	441
Tanvi Oka, Prachi Patankar, Shivani Rege, and Mrudul Dixit	
<b>Simple Square Slot Wearable Antenna with Coaxial Feed for ISM Band Applications at 5.3 GHz Frequency</b> .....	451
Vandana Yadav, Akhilesh Kumar, Prakhar Yadav, Amrees Pandey, and R. S. Yadav	
<b>Digital Storytelling and ICTs for Education to Foster Sustainable Development</b> .....	459
Komal Ramdey and Hasnain Bokhari	
<b>A Study on IoT-Based Health Monitoring Systems for Nurse Assistance</b> .....	467
Anju Pratap, Nithin Prince John, Greshma Johnson, Hanna Susan George, S. Devika, Anjal A. Jose, Kesia Sam, San Jos Varghese, Stevin Mathew Suban, and Salves Varghese	
<b>A Survey on Diabetes Mellitus Prediction Using Machine Learning Algorithms</b> .....	473
Rashmi Srivastava and Rajendra Kumar Dwivedi	

<b>DisasterNet: A Multi-label Disaster Aftermath Image Classification Model</b> .....	481
Subham Singh, Sandeep Ghosh, Amit Maity, Bikash Chandra Bag, Chaitali Koley, and Hirak Kumar Maity	
<b>Developing a Portrait Drawing Robot with FANUC LR Mate 200id/4s</b> .....	491
Ming Li	
<b>Comparative Analysis of Open Standards for Machine Learning Model Deployments</b> .....	499
Sachchidanand Singh, Naveen Singh, and Vinay Singh	
<b>Role of ICT in Emerging Towns and Cities</b> .....	509
S. Deepthi, R. Shanthi Priya, and S. RadhaKrishnan	
<b>A Web-Based Application for Data Collection and Report Generation Using Django</b> .....	519
C. Arun, S. Priyenka, L. Sagarika, and S. Sneha	
<b>On Trade Cloud Ecosystem Structure for Shared Learning</b> .....	529
Vaibhav E. Narawade and Vishal S. Badgujar	
<b>A Comparative Study of Local and Global Feature Detector and Descriptors for Image and Video Retrieval</b> .....	539
Ekta Sarda, Jayshree Jain, and Vaibhav Narawade	
<b>Extracting Attributes of Students Mental Health, Behaviour, Attendance and Performance in Academics During COVID-19 Pandemic using PCA Technique</b> .....	551
Vinayak Hegde, M. Shilpa, and M. S. Pallavi	
<b>Emotion and Gender Classification Using Convolution Neural Networks</b> .....	563
Chiranjeevi Pandi, K. Adi Narayana Reddy, Ramesh Alladi, V. Chandra Sekhar Reddy, and P. Sumithabhashini	
<b>Blockchain-Based Authentication Scheme to Secure Supermarket Transactions—A Survey</b> .....	575
L. Bilvashree, H. A. Brinda Nadig, K. R. Bharathi, B. S. Kavyashree, and K. Harshitha	
<b>Unified Flow Based Approach in Software-Defined Networks</b> .....	581
V. Deeban Chakravarthy, Dara Sujith Chandra, and Seemakurthi Sri Sathya Pavan	
<b>Towards a Framework for Context-Aware Intelligent Transportation System: Case of Kigali</b> .....	591
Gatera Antoine, Chomora Mikeka, Gaurav Bajpai, Andras Valko, and Kayalvizhi Jayavel	



<b>Part of Speech (PoS) Tagging for Konkani Language Using HMM</b> . . . . .	601
Annie Rajan and Ambuja Salgaonkar	
<b>Construction of Sports Articles Using Audio Commentary</b> . . . . .	611
Nilkamal More, Priya Fernandes, Nachiket Bhuta, Riya Suri, and Akshita Patil	
<b>Driver Drowsiness Detection Using Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and Driver Distraction Using Head Pose Estimation</b> . . . . .	619
T. V. N. S. R. Sri Mounika, P. H. Phanindra, N. V. V. N. Sai Charan, Y. Kranthi Kumar Reddy, and S. Govindu	
<b>Frequency Reconfigurable Patch Antenna for 5G Applications</b> . . . . .	629
Sanjeev Kumar, Sneha Prasad, Shubham Goswami, Tsarina Patnaik, and Satyaki Tatte	
<b>An Analysis of Future 5G Wireless Sensor Network</b> . . . . .	639
J. N. Swaminathan, M. Alagumeenaakshi, S. Umamaheswari, V. Jaikumar, and N. S. Kalyan Chakravathy	
<b>Building a Model of Signature Verification and Identification System Based on Deep CNN (ConvNet)</b> . . . . .	649
R. Shirisha, P. Divya, V. B. Aiswarya, S. Yoshitha, Jasim Sadique, and Anitha Premkumar	
<b>Product Review Analysis</b> . . . . .	659
L. V. Krishna Rao, S. Sai Vamsi, Sk. Sabiha Naazneen, S. Gowthami, and G. Bindu Bharati	
<b>Optimum Route Computation in a Chaotic Artificial Climbing Wall</b> . . .	671
Dibyendu Seal and Reshmita Seal	
<b>Integration of Attribute-Based Access Control in Microservices Architecture</b> . . . . .	681
Amandeep Singh, Vinay Raj, and Sadam Ravichandra	
<b>Analysis and Forecasting of Cybercrime Incident in India</b> . . . . .	691
Kousik Barik, Karabi Konar, Archita Banerjee, Saptarshi Das, and A. Abirami	
<b>Intelligent Indian Sign Language Recognition Systems: A Critical Review</b> . . . . .	703
Soumen Das, Saroj Kr. Biswas, Manomita Chakraborty, and Biswajit Purkayastha	
<b>Performance Analysis of Gujarati Script Recognition Using Multiclass and Multilabel Classification</b> . . . . .	715
Shweta Dalal, Sarosh Dastoor, Keyur Zaveri, Virti Choksi, and Karvin Shah	

<b>Real-Time Multi-Object Tracking of Pedestrians in a Video Using Convolution Neural Network and Deep SORT</b> .....	725
S. M. Praveenkumar, Prakashgouda Patil, and P. S. Hiremath	
<b>Student Engagement Awareness Dashboard for Asynchronous E-learning Environment</b> .....	737
Abdalganiy Kebede and Samit Bhattacharya	
<b>Customers' Willingness to Use New OTT Media Platforms in India</b> ....	751
Imon Chakraborty, Ravinder Kumar Verma, Mohit Hiteshkumar Joshi, and P. Vigneswara Ilavarasan	
<b>Application of Java Relationship Graphs to Academics for Detection of Plagiarism in Java Projects</b> .....	761
Divya Tyagi, Ritu Arora, and Yashvardhan Sharma	
<b>Predicting Appropriate Speed for Driving Based on External Factors</b> .....	773
Adwait Bhope, Aashay Zanpure, Vinod Kamat, Rohan Dhamale, Prashant Bhandari, Pruthi Joshi, Manisha Marathe, and Urmila Kalshetti	
<b>Real-Time Markerless Facial Landmark Detection Using Deep Learning</b> .....	785
Samyak Agarkar and Kapil Hande	
<b>A Novel Approach for Forecasting Account Receivables</b> .....	797
Parth Kapadia, Bintu Kadhiwala, Tejaswini Bahurupi, Het Dalal, Siddhi Jariwala, and Kshitij Naik	
<b>Power Supply Board and Switcher Board</b> .....	807
Yash Deep, Madan Chandrashekar, Arivalagan Venkatesan, Shiddalingeshwar Madli, K. A. Manjunath, and Surya Musunuri	
<b>Canonical Polyadic Decomposition of EEG Image Tensor for BCI Applications</b> .....	819
K. Keerthi Krishnan and K. P. Soman	
<b>Generalized Constraint Representation of Adjectives in PNL</b> .....	827
Bushra Siddique and M. M. Sufyan Beg	
<b>ECG Signal Compression Using Singular Value Decomposition and Implementation on PYNQ-Z2</b> .....	837
Mayur M. Sevak and Tanmay D. Pawar	
<b>Multi-Sectoral ICT Applications Toward Making Sustainable Rural India</b> .....	851
Sachi Gawai, Jyotshna Dongardive, and Pallavi Chatuphale	

**Implementation of e-Birth Registration Systems: Potential and Challenges ..... 861**  
Mahtab Shahin, Fatemeh Eskandari, Rozha Kamal Ahmed, and Dirk Draheim

**Risks Reduction of Rainfall-Induced Landslides-A Site-Specific Early Warning System (SSEWS) ..... 873**  
Martin Kuradusenge, Santhi Kumaran, Marco Zennaro, and Jean Baptiste Minani

**Decoding Visual Covert Attention Shift from EEG for Use in BCI ..... 883**  
Swati Aggarwal, Nupur Chugh, and Arnav Balyan

**Soft Embedded Cores for FPGA ..... 895**  
Praveen Kalkundri, Hansraj Guhilot, and Kalkundri Ravi

**Social Media Enabled Rehabilitation Services: Influence of Covid-19 Metaphors ..... 907**  
Sreejith Alathur

**Fuzzy Logic Based Support Vector Regression (SVR) Model for Software Cost Estimation Using Machine Learning ..... 917**  
Kamal Upreti, Umesh Kumar Singh, Rituraj Jain, Karamjeet Kaur, and A. K. Sharma

**Author Index ..... 929**

# Editors and Contributors

## About the Editors

**Milan Tuba** received his B.S. and M.S. in Mathematics, and M.S., M.Ph., and Ph.D. in Computer Science from the University of Belgrade and New York University, Courant Institute of Mathematical Sciences. He is a full professor and the Vice Rector for International Relations, Singidunum University, Belgrade, Serbia and was the Head of the Department for Mathematical Sciences at State University of Novi Pazar and the Dean of the Graduate School of Computer Science at John Naisbitt University. He is listed in both lists of the 2020 World's Top 2% Scientists in all disciplines by Stanford University, the one for overall career contribution and the second for contribution in the previous year. He was invited and delivered around 60 keynote and plenary lectures at international congresses and conferences. Professor Tuba is the author or coauthor of almost 300 scientific papers. His research interest includes nature-inspired optimizations applied to image processing, computer networks, and neural networks. He is an editor, coeditor or member of the editorial board or scientific committee of a number of scientific journals and conferences. Member of the ACM, IEEE, AMS, SIAM, IFNA, IASEI.

**Shyam Akashe** is a Professor at ITM University, Gwalior, Madhya Pradesh, India. He completed his Ph.D. at Thapar University, Punjab, and his M.Tech. in Electronics and Communication Engineering at the Institute of Technology & Management, Gwalior. He has authored 190 publications, including more than 50 papers in SCI-indexed journals. His main research focus is low-power system on chip (SoC) applications in which static random access memories (SRAMs) are omnipresent. He has authored two books entitled Moore's Law Alive: Gate-All-Around (GAA) Next Generation Transistor published by LAMBERT Academic Publishing, Germany, and Low Power High Speed CMOS Multiplexer Design published by Nova Science Publishers, Inc., New York, USA. He has also published over 120 papers in leading national and international refereed journals of repute. Dr. Akashe has participated in numerous national and international conferences and presented over 100 papers.

**Amit Joshi** is currently the Director of Global Knowledge Research Foundation, also an Entrepreneur & Researcher who has completed his Masters and research in the areas of cloud computing and cryptography in medical imaging. Dr. Joshi has an experience of around 10 years in academic and industry in prestigious organizations. Dr. Joshi is an active member of ACM, IEEE, CSI, AMIE, IACSIT, Singapore, IDES, ACEEE, NPA, and many other professional societies. Currently, Dr. Joshi is the International Chair of InterYIT at International Federation of Information Processing (IFIP, Austria). He has presented and published more than 50 papers in national and international journals/conferences of IEEE and ACM. Dr. Joshi has also edited more than 40 books which are published by Springer, ACM, and other reputed publishers. Dr. Joshi has also organized more than 50 national and international conferences and programs in association with ACM, Springer, and IEEE to name a few across different countries including India, UK, Europe, USA, Canada, Thailand, Egypt, and many more.

## Contributors

**A. R. Abdul Rajak** EEE Department, Birla Institute of Technology and Science, Dubai, United Arab Emirates

**Manish Kumar Abhishek** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

**A. Abirami** Bannari Amman Institute of Technology, Erode, India

**K. Adi Narayana Reddy** BVRIT Hyderabad College of Engineering for Women, Hyderabad, India

**S. Aditya** CSE, Thiagarajar College of Engineering, Madurai, India

**M. Afshar Alam** Department of Computer Science and Engineering, School of Engineering Sciences and Technology, New Delhi, India

**Samyak Agarkar** CSE Department, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

**Ayush Agarwal** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Parul Agarwal** Department of Computer Science and Engineering, New Delhi, Delhi, India

**Swati Aggarwal** Netaji Subhas Institute of Technology, New Delhi, India

**Pallavi Ahire** Department of Technology, Savitribai Phule Pune University, Pune, India

**Jawed Ahmed** Department of Computer Science and Engineering, School of Engineering Sciences and Technology, New Delhi, Delhi, India

**Rozha Kamal Ahmed** Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

**R. Aishwarya** Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India

**V. B. Aiswarya** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**M. Ajitha** Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India

**M. Alagumeenaakshi** Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India

**Sreejith Alathur** National Institute of Technology Karnataka, Mangalore, India

**Ramesh Alladi** ACE Engineering College, Hyderabad, India

**Shreya Anekar** Department of Information Technology, MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

**Gatera Antoine** African Center of Excellence in Internet of Things, College of Science and Technology, University of Rwanda, Kigali, Rwanda

**Ritu Arora** Birla Institute of Technology and Science, Pilani, India

**Shweta Arora** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**C. Arun** Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Chennai, India

**Vishal S. Badgujar** Department of Information Technology, A. P. Shah Institute of Technology, Thane, Mumbai, India

**Mahima Badjate** MKSSS's Cummins College of Engineering for Women, Pune, India

**Bikash Chandra Bag** National Institute of Technology Mizoram, Chaltlang, Aizawl, India

**Tejaswini Bahurupi** Computer Engineering Department, Sarvajanic College of Engineering and Technology, Surat, India

**Gaurav Bajpai** Department of Computer and Software Engineering, College of Science and Technology, University of Rwanda, Kigali, Rwanda

**Soham Balkothe** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Arnav Balyan** Netaji Subhas Institute of Technology, New Delhi, India

**Archita Banerjee** JIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

**Kousik Barik** JIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

**Pranav Bathija** Thadomal Shahani Engineering College, Mumbai, India

**Devyani Bhamare** Sanjivani College of Engineering, Kopergaon, India

**Prashant Bhandari** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**K. R. Bharathi** Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Ashish Bhat** Thadomal Shahani Engineering College, Mumbai, India

**Samit Bhattacharya** Indian Institute of Technology, Guwahati, Assam, India

**Adwait Bhope** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**Nachiket Bhuta** Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India

**M. Bhuvaneshwari** CSE, Thiagarajar College of Engineering, Madurai, India

**L. Bilvashree** Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**G. Bindu Bharati** JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Narendra Bisht** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Saroj Kr. Biswas** NIT Silchar, Fakiratilla, Silchar, Assam, India

**Hasnain Bokhari** University of Erfurt, Erfurt, Germany

**Jyoti G. Borade** Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, India

**H. A. Brinda Nadig** Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**A. K. Buvanagri** Department of Information Technology, SSN College of Engineering, Chennai, Tamil Nadu, India

**Imon Chakraborty** UQ-IITD Academy of Research, Indian Institute of Technology, Delhi, India

**Manomita Chakraborty** NIT Silchar, Fakiratilla, Silchar, Assam, India

**V. Deeban Chakravarthy** SRM University, Kattankulathur, Chennai, Tamilnadu, India

**S. H. Chandak** Department of IT, PICT, Pune, India

**Dara Sujith Chandra** SRM University, Kattankulathur, Chennai, Tamilnadu, India

**V. Chandra Sekhar Reddy** ACE Engineering College, Hyderabad, India

**Madan Chandrashekar** Infineon Technologies India Private Limited, Bengaluru, India

**Soumyadip Chatterjee** Edifixio India Private Limited, Newtown, Kolkata, West Bengal, India

**Pallavi Chatuphale** University Department of Computer Sciences, University of Mumbai, Mumbai, India

**Harsh Chawla** Thadomal Shahani Engineering College, Mumbai, India

**R. R. Chhajed** Department of IT, PICT, Pune, India

**Dhairya Chheda** Shah and Anchor Kutchhi Engineering College, Mumbai, India

**Virti Choksi** Sarvajani College of Engineering and Technology, Gujarat, India

**Nupur Chugh** Netaji Subhas Institute of Technology, New Delhi, India

**John Colaco** Department of Electronics and Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

**Het Dalal** Computer Engineering Department, Sarvajani College of Engineering and Technology, Surat, India

**Shweta Dalal** Sarvajani College of Engineering and Technology, Gujarat, India

**Dhara N. Darji** Department of Computer Science, Ganpat University, Gujarat, India

**L. N. Das** Department of Applied Mathematics, Delhi Technological University, New Delhi, India

**Saptarshi Das** JIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

**Soumen Das** NIT Silchar, Fakiratilla, Silchar, Assam, India

**Sarosh Dastoor** Sarvajani College of Engineering and Technology, Gujarat, India

**Yash Deep** Infineon Technologies India Private Limited, Bengaluru, India

**S. Deepthi** Prime Nest College of Architecture and Planning, Trichy, Affiliated To Anna University, Chennai, Tamilnadu, India

**Jaimin Desai** NMIMS University, Mumbai, India

**A. V. Deshpande** Department of Computer Engineering, S.K.N. College of Engineering, Pune, India



**Arti Deshpande** Thadomal Shahani Engineering College, Mumbai, India

**Prajakta Deshpande** MKSSS's Cummins College of Engineering for Women, Pune, India

**S. Devika** Saintgits College of Engineering, Kottayam, India

**Rohan Dhamale** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**S. C. Dharmadhikari** Department of IT, PICT, Pune, India

**Roshani K. Dharme** G.H. Raison Institute of Engineering and Technology, Nagpur, India

**P. Divya** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**Mrudul Dixit** MKSSS'S Cummins College of Engineering for Women, Pune, India

**Jyotshna Dongardive** University Department of Computer Sciences, University of Mumbai, Mumbai, India

**Dirk Draheim** Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

**Asha Durafe** Shah & Anchor Kutchhi Engineering College, Chembur, Mumbai, India;

Sir Padampat Singhania University, Udaipur, Rajasthan, India

**Rajendra Kumar Dwivedi** Department of IT, CAMMMUT, Gorakhpur, UP, India

**Fatemeh Eskandari** Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

**Priya Fernandes** Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India

**M. Ganeshan** Department of Computer Science and Engineering, Jain Deemed to be University, Bangalore, India

**Sachi Gawai** Ram Rao Adik Institute of Technology, University of Mumbai, Mumbai, India

**Ujwala Gawande** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Sakshi Gehani** Narsee Monjee Institute of Management Studies, Mumbai, India

**Hanna Susan George** Saintgits College of Engineering, Kottayam, India

**D. I. George Amalarethinam** Department of Computer Science, Jamal Mohamed College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

**Sumedh Ghatwai** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Ramyani Ghosh** PES University, Bangalore, Karnataka, India

**Sandeep Ghosh** College of Engineering and Management, Kolaghat, West Bengal, India

**Shubham Goswami** Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India

**S. Govindu** Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India

**S. Gowthami** JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Hansraj Guhilot** K. C. College of Engineering and Management Studies, Mumbai, India

**Kapil Hande** CSE Department, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

**M. Haripriya** CSE, Thiagarajar College of Engineering, Madurai, India

**K. Harshitha** Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Vinayak Hegde** Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Mysuru, India

**P. S. Hiremath** KLE Technological University, Hubballi, Karnataka, India

**Shivanand C. Hiremath** R. N. Shetty Polytechnic College, Belagavi, Karnataka, India

**P. Vigneswara Ilavarasan** Department of Management Studies, Indian Institute of Technology, Delhi, India

**Purva Ingle** MIT World Peace University, Pune, India

**V. Jaikumar** QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India

**Jayshree Jain** Department of Computer Engineering, Pacific Academic Higher Education and Research University, Udaipur, India

**Rituraj Jain** Department of Electrical and Computer Engineering, Wollega University, Nekemte, Ethiopia

**Siddhi Jariwala** Computer Engineering Department, Sarvajanik College of Engineering and Technology, Surat, India

**Kayalvizhi Jayavel** Department of Information Technology, SRM Institute of Science and Technology, Chennai, India

**Ravi Shankar Jha** KIIT School of Management, KIIT University (Institution of Eminence), Bhubaneswar, India

**Nithin Prince John** Saintgits College of Engineering, Kottayam, India

**Greshma Johnson** Saintgits College of Engineering, Kottayam, India

**Anjal A. Jose** Saintgits College of Engineering, Kottayam, India

**Kavita Ajay Joshi** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Mohit Hiteshkumar Joshi** Department of Management Studies, Indian Institute of Technology, Delhi, India

**Pratyush Joshi** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Prutha Joshi** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**Bintu Kadhiwala** Computer Engineering Department, Sarvajanik College of Engineering and Technology, Surat, India

**Sumit Kale** Sinhgad Institute of Technology, Lonavla, India

**Praveen Kalkundri** Gogte Institute of Technology, Udyambag, Belgaum, Karnataka, India

**Urmila Kalshetti** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**N. S. Kalyan Chakravathy** QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India

**Vinod Kamat** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**Rutuja Kamble** Sinhgad Institute of Technology, Lonavla, India

**Parth Kapadia** Computer Engineering Department, Sarvajanik College of Engineering and Technology, Surat, India

**S. Karthika** Department of Information Technology, SSN College of Engineering, Chennai, Tamil Nadu, India

**Karamjeet Kaur** Department of Computer Science, MGM College of Engineering and Technology, Noida, India

**Shubhpreet Kaur** Lovely Professional University, Punjab, India

**Tarandeep Kaur** Lovely Professional University, Phagwara, Punjab, India

**B. S. Kavyashree** Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Abdalganiy Kebede** Indian Institute of Technology, Guwahati, Assam, India

**Amit Kharat** DeepTek Inc, Pune, India

**Arvind W. Kiwelekar** Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, India

**Chaitali Koley** National Institute of Technology Mizoram, Chaltlang, Aizawl, India

**Shivani Krishna Kompella** Birla Institute of Technology and Science, Dubai, United Arab Emirates

**Karabi Konar** JIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

**Pranali Kosamkar** MIT World Peace University, Pune, India

**Y. Kranthi Kumar Reddy** Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Krishna, Andhra Pradesh, India

**L. V. Krishna Rao** JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**K. Keerthi Krishnan** Center for Computational Engineering Networking (CEN), Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Manish Gawali Viraj Kulkarni** Deeptek Inc., Pune, India

**Viraj Kulkarni** DeepTek Inc, Pune, India

**Vrushali Kulkarni** MIT World Peace University, Pune, India

**Akhilesh Kumar** Electronics Engineering Department, Institute of Engineering and Rural Technology, Prayagraj, UP, India

**N. S. Kumar** PES University, Bangalore, Karnataka, India

**Sanjeev Kumar** Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India

**Santhi Kumaran** School of ICT, The Copperbelt University, Kitwe, Zambia

**Harsha C. Kunwar** G.H. Raison Institute of Engineering and Technology, Nagpur, India

**Martin Kuradusenge** African Centre of Excellence in Internet of Things, University of Rwanda, Kigali, Rwanda

**Ming Li** University of Michigan-Flint, Flint, MI, USA

**Rajesh B. Lohani** Department of Electronics and Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

**Rohit Lokwani** DeepTek Inc, Pune, India

**Harshit Madaan** Indian Institute of Science Education and Research, Pune, India; DeepTek Inc., Pune, India

**J. Madhupriya** Department of Computer Science, Jamal Mohamed College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

**Shiddalingeshwar Madli** Infineon Technologies India Private Limited, Bengaluru, India

**Parikshit N. Mahalle** Department of Computer Engineering, S.K.N. College of Engineering, Pune, India

**Geetanjali S. Mahamunkar** Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, India

**Amit Maity** College of Engineering and Management, Kolaghat, West Bengal, India

**Hirak Kumar Maity** College of Engineering and Management, Kolaghat, West Bengal, India

**Jayashree D. Mallapur** Basaveshwar Engineering College, Bagalkot, Karnataka, India

**Surajit Mandal** B.P. Poddar Institute of Management and Technology, Kolkata, West Bengal, India

**V. V. Mandhare** Sanjivani College of Engineering, Kopargaon, India; Pravara Rural Engineering College, Loni, India

**K. A. Manjunath** Infineon Technologies India Private Limited, Bengaluru, India

**Manisha Marathe** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**R. Meenakshi** Department of Information Technology, SSN College of Engineering, Chennai, Tamil Nadu, India

**V. S. Meghana** PES University, Bangalore, Karnataka, India

**Nalini Mhetre** Department of Computer Engineering, Sinhgad College of Engineering, Pune, India

**Chomora Mikeka** Department of Physics, University of Malawi, Zomba, Malawi

**Jean Baptiste Minani** Concordia University, Montreal, QC, Canada

**Vedant Modi** MIT World Peace University, Pune, India

**Monu** Department of Applied Mathematics, Delhi Technological University, New Delhi, India

**Nilkamal More** Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India

**Surya Musunuri** Infineon Technologies India Private Limited, Bengaluru, India

**Rohit Mutalik** Shah and Anchor Kutchhi Engineering College, Mumbai, India

**Kshittij Naik** Computer Engineering Department, Sarvajanic College of Engineering and Technology, Surat, India

**Sapan Naik** Babu Madhav Institute of Information Technology, Uka Tarsadia University, Surat, Gujarat, India

**Deepa Nainwal** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Girish Nainwal** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Vaibhav Narawade** Department of Computer Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, India

**Vaibhav E. Narawade** Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India

**Preet Nain Kaur Narula** Graphic Era Hill University, Bhimtal, Uttarakhand, India

**Shreyas Nemade** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Laxman D. Netak** Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, India

**Tanvi Oka** MKSSS'S Cummins College of Engineering for Women, Pune, India

**Sushila Paliwal** Department of Computer Science and Engineering, School of Engineering Sciences and Technology, New Delhi, India

**M. S. Pallavi** Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Mysuru, India

**Yogesh H. Palve** MVPS's KBT COE, Nashik, India

**Siba Panda** Mentor, Narsee Monjee Institute of Management Studies, Mumbai, India

**Amrees Pandey** Department of Electronics and Communication, University of Allahabad, Prayagraj, UP, India

**Chiranjeevi Pandi** ACE Engineering College, Hyderabad, India

**Aniruddha Pant** Deeptek Inc., Pune, India

**Satyen M. Parikh** Faculty of Computer Applications, Ganpat University, Gujarat, India

**Suraiya Parveen** Department of Computer Science and Engineering, School of Engineering Sciences and Technology, New Delhi, India

**Sagupha Parween** College of Engineering and Technology, Bhubaneswar, Odisha, India

**Prachi Patankar** MKSSS'S Cummins College of Engineering for Women, Pune, India

**Hiral R. Patel** Department of Computer Science, Ganpat University, Gujarat, India

**Toral Patel** MKSSS's Cummins College of Engineering for Women, Pune, India

**Vinod Patidar** Sir Padampat Singhania University, Udaipur, Rajasthan, India

**Akshita Patil** Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India

**Prakashgouda Patil** KLE Technological University, Hubballi, Karnataka, India

**Tsarina Patnaik** Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India

**Seemakurthi Sri Sathya Pavan** SRM University, Kattankulathur, Chennai, Tamilnadu, India

**Tanmay D. Pawar** Electronics Engineering, Birla Vishvakarma Mahavidyalaya, Anand, Gujarat, India

**Yash Pazare** Sinhgad Institute of Technology, Lonavla, India

**P. H. Phanindra** Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Krishna, Andhra Pradesh, India

**Saniya Ponkshe** MKSSS's Cummins College of Engineering for Women, Pune, India

**Sneha Prasad** Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India

**Anju Pratap** Saintgits College of Engineering, Kottayam, India

**S. M. Praveenkumar** KLE Technological University, Hubballi, Karnataka, India

**Anitha Premkumar** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**S. Priyenka** Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Chennai, India

**Biswajit Purkayastha** NIT Silchar, Fakiratilla, Silchar, Assam, India

**S. RadhaKrishnan** Prime Nest College of Architecture and Planning, Trichy, Affiliated To Anna University, Chennai, Tamilnadu, India

**Vinay Raj** BVRIT HYDERABAD College of Engineering for Women, Hyderabad, India

**Annie Rajan** Dhempe College of Arts and Science, Panaji, India

**A. Rajesh** Department of Computer Science and Engineering, Jain Deemed to be University, Bangalore, India

**Varshini Rajesh** Birla Institute of Technology and Science, Dubai, United Arab Emirates

**D. Rajeswara Rao** Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

**Ashish P. Ramdasi** Hindustan Institute of Technology and Science, Chennai, India

**Komal Ramdey** Willy Brandt School of Public Policy, University of Erfurt, Erfurt, Germany

**Rasya Ramesh** PES University, Bangalore, Karnataka, India

**Yash Rana** Babu Madhav Institute of Information Technology, Uka Tarsadia University, Surat, Gujarat, India

**Kalkundri Ravi** K. L. S. Gogte Institute of Technology, Belagavi, Karnataka, India

**Sadam Ravichandra** National Institute of Technology, Warangal, India

**Shivani Rege** MKSSS'S Cummins College of Engineering for Women, Pune, India

**Sneha Rohida** MKSSS's Cummins College of Engineering for Women, Pune, India

**Sk. Sabiha Naazneen** JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Jasim Sadique** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**L. Sagarika** Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Chennai, India

**Priti Ranjan Sahoo** KIIT School of Management, KIIT University (Institution of Eminence), Bhubaneswar, India

**N. V. V. N. Sai Charan** Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Krishna, Andhra Pradesh, India



**S. Sai Vamsi** JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Saziabegum Saiyed** S.S.Agrawal Institute of Management and Technology, Navsari, Gujarat, India

**Priti Sajja** G.H.Patel PG Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidhyanagar, Anand, Gujarat, India

**Ambuja Salgaonkar** Department of Computer Science, University of Mumbai, Mumbai, India

**Kesia Sam** Saintgits College of Engineering, Kottayam, India

**Ekta Sarda** Faculty of Computer Engineering, Pacific Academy of Higher Education and Research University, Udaipur, India

**R. Sarunisha** CSE, Thiagarajar College of Engineering, Madurai, India

**S. Sathyalakshmi** Hindustan Institute of Technology and Science, Chennai, India

**Dibyendu Seal** Tata Consultancy Services, Kolkata, India

**Reshmita Seal** Techno Main Saltlake, Kolkata, India

**Jitesh Seth** Indian Institute of Science Education and Research, Pune, India; DeepTek Inc, Pune, India

**Mayur M. Sevak** Gujarat Technological University, Gujarat, India

**Karvin Shah** Sarvajanic College of Engineering and Technology, Gujarat, India

**Mahtab Shahin** Information Systems Group, Tallinn University of Technology, Tallinn, Estonia

**Zeeshan Shaikh** Shah and Anchor Kutchhi Engineering College, Mumbai, India

**R. Shanthi Priya** School of Architecture and Interior Design, SRM Institute of Science and Technology, Kattankulanthur, Tamilnadu, India

**A. K. Sharma** Department of CSI, University of Kota, Kota, India

**Ashok Sharma** Lovely Professional University, Punjab, India

**Yashvardhan Sharma** Birla Institute of Technology and Science, Pilani, India

**A. Sheryl Oliver** Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai, India

**Shilpa** Lovely Professional University, Phagwara, Punjab, India

**M. Shilpa** Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Mysuru, India

**R. Shirisha** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**Nishant Shrikhande** Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

**Bushra Siddique** Aligarh Muslim University, Aligarh, India

**Amandeep Singh** National Institute of Technology, Warangal, India

**Chanavi Singh** MIT World Peace University, Pune, India

**Naveen Singh** Tata Consultancy Services (TCS), Nagpur, India

**Sachchidanand Singh** IBM India Software Labs (ISL), Pune, India

**Subham Singh** College of Engineering and Management, Kolaghat, West Bengal, India

**Umesh Kumar Singh** Institute of Computer Science, Vikram University, Ujjain, India

**Vinay Singh** Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur, India

**Smrutirekha** KIIT School of Management, KIIT University (Institution of Eminence), Bhubaneswar, India

**S. Sneha** Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Chennai, India

**K. P. Soman** Center for Computational Engineering Networking (CEN), Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

**Harsha Sonune** Department of Information Technology, MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

**T. V. N. S. R. Sri Mounika** Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Krishna, Andhra Pradesh, India

**Rashmi Srivastava** Department of IT, CA MMMUT, Gorakhpur, UP, India

**Stevin Mathew Suban** Saintgits College of Engineering, Kottayam, India

**M. M. Sufyan Beg** Aligarh Muslim University, Aligarh, India

**P. Sumithabhashini** SAMSKRUTI College of Engineering and Technology, Hyderabad, India

**Riya Suri** Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India

**Jaya R. Surywanshi** MVPS's KBT COE, Nashik, India

**J. N. Swaminathan** QIS College of Engineering and Technology, Ongole, Andhra Pradesh, India

**Mona Tatari** Department of Computer Science and Engineering, New Delhi, Delhi, India

**Satyaki Tatte** Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India

**Viraj Thakkar** Babu Madhav Institute of Information Technology, Uka Tarsadia University, Surat, Gujrat, India

**Dhanashree Toradmalle** Shah and Anchor Kutchhi Engineering College, Mumbai, India

**Aruna Tripathy** College of Engineering and Technology, Bhubaneswar, Odisha, India

**Divya Tyagi** Birla Institute of Technology and Science, Pilani, India

**S. Umamaheswari** Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India

**Kamal Upreti** Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology & Management, Delhi, India

**Andras Valko** African Center of Excellence in Internet of Things, College of Science and Technology, University of Rwanda, Kigali, Rwanda

**Salves Varghese** Saintgits College of Engineering, Kottayam, India

**San Jos Varghese** Saintgits College of Engineering, Kottayam, India

**Arivalagan Venkatesan** Infineon Technologies India Private Limited, Bengaluru, India

**Ravinder Kumar Verma** Department of Management Studies, Indian Institute of Technology, Delhi, India

**V. Vignaraj Ananth** CSE, Thiagarajar College of Engineering, Madurai, India

**Kashika Wanchoo** NMIMS University, Mumbai, India

**Krishna Yadav** Sinhgad Institute of Technology, Lonavla, India

**Prakhar Yadav** Department of Electronics and Communication, University of Allahabad, Prayagraj, UP, India

**R. S. Yadav** Department of Electronics and Communication, University of Allahabad, Prayagraj, UP, India

**Vandana Yadav** Department of Electronics and Communication, University of Allahabad, Prayagraj, UP, India

**Samrudhi Yeginwar** Department of Information Technology, MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

**S. Yoshitha** Department of Computer Science, Presidency University, Bengaluru, Karnataka, India

**Aashay Zangpure** Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India

**Keyur Zaveri** Sarvajanik College of Engineering and Technology, Gujarat, India

**Marco Zennaro** International Centre of Theoretical Physics, Strada Costiera, Trieste, Italy

# Literature Survey for MIMO Systems and Interference Cancellation in MIMO



Varshini Rajesh, Shivani Krishna Kompella, and A. R. Abdul Rajak

**Abstract** This survey paper discusses a reliable technique, applicable for extensive spectral efficiency, MIMO technologies. In recent times, MIMO systems and their various types have attracted tremendous analysis and research in wireless communications. Although we are blessed to be introduced to the several advantages of this technology, it faces so many issues and must surmount them to emerge successful in the field of transceiver communication. One of the pivotal challenges is the quality of reception at the base station because its difficulty exponentially increases with a massive variety of antennas. The key factors to be considered during a wireless network design are channel capacity and spectral efficiency. Intra-cell interference plays a major role in determining the increase in spectral efficiency of large power transmits. In cellular networks, the per-cell spectral efficiency is mostly challenging to improve due to the interference. This paper analyses performance and concomitance of MIMO radars and MIMO communication systems and different approaches to perform interference cancellation in MIMO, MIMO-OFDM, multi-user MIMO antenna networks and detectors, massive MIMO, OAM-MIMO, and full-duplex MIMO systems.

**Keywords** Wireless communications · MIMO radars and systems · Massive MIMO · Passive and successive interference cancellation · Analog cancellation

---

V. Rajesh (✉) · S. K. Kompella · A. R. Abdul Rajak  
Birla Institute of Technology and Science, Pilani, Dubai, United Arab Emirates  
e-mail: [f20170237@dubai.bits-pilani.ac.in](mailto:f20170237@dubai.bits-pilani.ac.in)

S. K. Kompella  
e-mail: [f20170243@dubai.bits-pilani.ac.in](mailto:f20170243@dubai.bits-pilani.ac.in)

A. R. Abdul Rajak  
e-mail: [abdulrazak@dubai.bits-pilani.ac.in](mailto:abdulrazak@dubai.bits-pilani.ac.in)

# 1 Introduction

## 1.1 A Description of MIMO Systems

In the last two to three decades, wireless and mobile telecommunication networks have experienced enormous growth opportunities to emerge as a successful field. The rapidly increasing demand for higher data rates, higher capacity to provide broadband access to indoor and outdoor regions in dense urban areas, and increased user agility are some of the factors responsible in enabling reliable transmissions over mobile radio channels. Multiple-input multiple-output (MIMO) is a wireless antenna technology for communications which aims to incorporate multiple transmitters and receivers and transfer more data while minimizing errors. The theoretical concept of MIMO has been around for the past few years, but its practical application in wireless communications is more novel and are used in technologies like Wi-Fi and LTE nowadays [1].

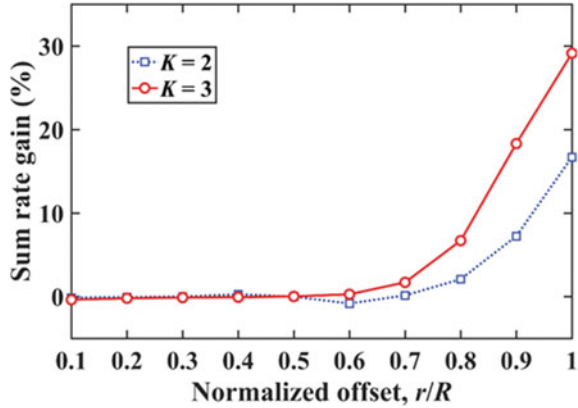
# 2 A Study on MIMO Systems

Firstly, calculation and study of two-user indoor 3D-MIMO channel in an office scenario were conducted via the measurement campaign using a vector network analyzer (VNA) as it offers a huge MIMO measurement bandwidth. The 3D-MIMO channel mm-wave is a fast, 3D inverse Fourier transform (IFFT) characterization inside the wavenumber and time domain [2]. In the approximate resolution of multipath elements, the 3D-MIMO is shown to be enormous and considerably higher than non-massive 3D-MIMO. It is observed that a  $16 \times 16$  URA has better multipath components resolution than what a  $4 \times 4$  URA would have.

A novel normalized gain difference power allocation (NGDPA) approach for the exploitation of channel conditions was suggested in this paper for the adaptive and effective power assignment of low computational complexities in indoor MIMO-NOMA-VLC (visible light communication) systems [3]. The results indicate that NOMA with NGDPA in an indoor  $2 \times 2$  MIMO VLC system is much better than NOMA with GRPA [4]. In contrast, the sum rate benefit of NGDPA in comparison with GRPA, as shown in Fig. 1, is considerably greater for many more customers in the system. MIMO-NOMA, with the NGDPA solution suggested, is therefore favorable for potential multi-user VLC systems with high speed.

As seen in the simulation experiments, the method offers MIMO radar more freedom and thus improved radar efficiency than previous works that focused on transmitting communication symbols to synthesize a pattern of radar beams [5]. Simulation studies have shown that the efficiency of the radar beam patterns and estimation of angle achieved with the dual-functions method proposed is analogous with the optimum radar-only method.

**Fig. 1** Sum rate gain of NGDPA over GRPA in NOMA versus normalized offset



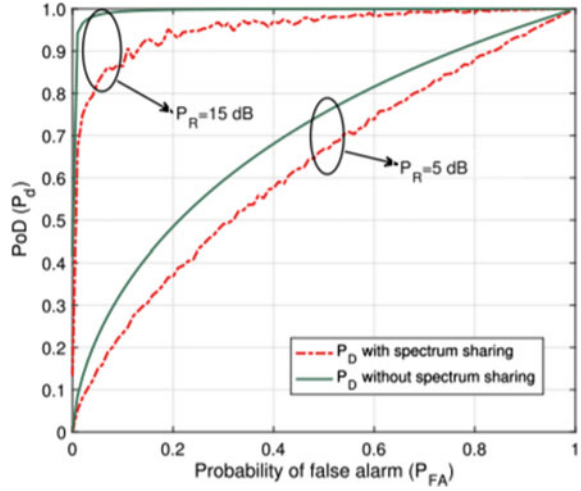
Complexity has been minimized by lowering the number of iterations in the adaptive method and the preamble size used. Furthermore, the IASCE algorithm exceeds the computer complexity of the LCSCE algorithm. The two suggested algorithms not only minimize the complexity, they save the bandwidth by integrating short preamble lengths [6].

Spatial detection algorithms like ZF and MMSE suboptimal spatial multiplexing detection algorithms are presented. The spatial association between the MIMO channel coefficients induces the loss of channel capacity and affects the efficiency of the bit rate error. Measured channel coefficients that are used to calculate the correlation matrices are used as the conditions for MIMO research modes [7]. Mathematical simulation reveals that a spatial association affects the ZF's BER performance and the MMSE detection algorithms. Although a joint transceiver architecture strategy is suggested in the CS BS and users, in compliance with users' quality of service (QoS) restrictions and CS transmission capacity, to increase the likelihood of identification (PoD) of the MIMO RS, null-space waveform space can be utilized to minimize interference from RS to CS, also observed in Fig. 2 [8]. The technique suggested [9], in particular, RS PoD's efficiency, optimizing its lower limit, derived from the monotonously increasing relation between PoD and its parameter of non-centeredness that the structure is generally intended to ensure that cellular networks of the future are established effectively in collaboration with federal incumbents who can function within the same scope of capital.

A structure consisting of the key iterative algorithm through two sub-algorithms comprised in its procedure. The primary sub-algorithm is expended to optimize the signal-to-interference-plus-noise ratio of the radar system, while the secondary is utilized to optimize the mean rate of the communication system. The key iterative algorithm is recommended to refine the propagation cooperatively and to obtain radiated filters of the two systems by rotating the optimization of the two sub-algorithms. Numerical checking of the convergence of the algorithm and its sub-algorithms [10].

A combined optimization of the covariance matrix and the transmit radar waveform contact is intended for optimizing the rate of contact, with restrictions about the

**Fig. 2** Graph showing probability of detection at MIMO RS versus probability of false alarm at MIMO RS



resemblance with the normal radar waveform and radar and communication energies. MIMO links are aimed to achieve interference cancellation, which is explored in the next pages, by ground-breaking technologies and sophisticated signal processing algorithms [11].

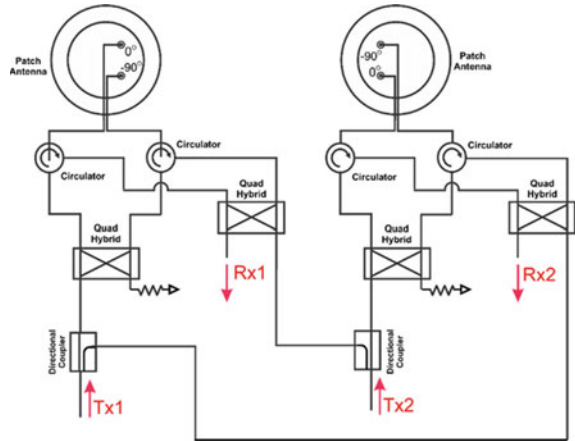
### 3 A Study on Interference Cancellation in MIMO Systems

Self-interference and cross-interference passive cancellation when a  $2 \times 2$  MIMO node is tested in STAR (simultaneous transmit and receive) operation has been analyzed [12]. Circularly polarized antennas help in cross-cancellation at coupling transceiver antennas, and further improvisation can be achieved by a feedforward path among transceivers. A self-interference cancellation of total value more than 40 dB and a value of above 46 dB for cross-interference cancellation can be observed and measured in a system with two antennas connected in a balanced manner of feed network. A side-by-side antenna pair to improve cross-interference cancellation has been shown in Fig. 3.

Hence, it has been concluded that in a node of the form  $2 \times 2$  STAR, self- and cross-interference cancellation is efficient by using passive cancellation as well as a balanced feed network. It has also been noted that simplification of design occurs while accomplishing greater levels of self- and cross-interference cancellation at the transmitter and receiver antenna in an absolute  $2 \times 2$  STAR MIMO system. The authors [13] propose an algorithm for interference cancellation units (ICUs) to relieve hardware complexity by an optimized, pipelined, and parallel construction in MIMO detectors. Making use of this proposed architecture in place of its standard counterpart, eviction of 3.5 times more adders is proved to be possible in the detector



**Fig. 3** Side-by-side antenna pair for feed forward mechanism



for several MIMO alignments. Complexity was reduced, and the architecture was indeed confirmed effective by using a prototype of a MIMO detector which processed 7.15 Gbps while expending 39.9 pJ/b. In conclusion, ICUs were redesigned in order to maximize the reduction of complexity without drawing out the critical path.

Device-to-device (D2D) communication involving multi-user multi-input multi-output (MU-MIMO) cellular communication networks is investigated. Certain assumptions are made before evaluating the given procedure [14], such that down-link time–frequency resources of cellular links are being reused by D2D users and multiple antennas are being equipped at the base station (BS). When interference cancellation (IC) pre-coding strategy is deployed at the BS, the ergodic realizable sum rate of a system is being researched. For each link, the received signal-to-interference-noise ratio (SINR) distributions and deriving an exact ergodic achievable aggregate rate of the whole system including closed-form expressions are looked into. The count of antennas at BS and the network topology is presented to be arbitrary for all the expressions.

A technique known as ordering sequential interference cancellation (OSIC) detection is applied for estimating soft-output minimum mean square error (MMSE) in channel coded wireless MIMO systems. At the cost of minimal complexity increase, simulation results of this proposed novel detector are commendable because they help in the reduction of residual error floor caused by channel estimation errors, in comparison with the prevailing soft-output MMSE MIMO detector. With modulation of lower order and systems with high SNR, the operation of the aimed detector is observed to be enhanced than the aforementioned strategies.

For systems with defective channel estimation, vigorous SC-FDE MIMO design is being investigated. Channel estimation linked with TDM training vectors and matrices that are corresponding are represented by Subscript “ $T_r$ ”.  $\mathcal{X}_j$  represents an elementary matrix that is a Toeplitz matrix with dimensions N-by-L and can be shown as denoted in Eq. 1 [15]:

$$\mathcal{X}_j = \begin{bmatrix} x_{0,j}; T_r & x_{N-1,j}; T_r & \dots & x_{N-L+1,j}; T_r \\ x_{1,j}; T_r & x_{0,j}; T_r & \dots & x_{N-L+2,j}; T_r \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-1,j}; T_r & x_{0,j}; T_r & \dots & x_{N-L,j}; T_r \end{bmatrix} \quad (1)$$

With the main diagonal elements as zero,  $c_0$  can be defined as a lower triangular matrix,  $b_n$  can be given by Eq. 2 [15] as follows,

$$b_n = \sum_{m=0}^B c_m d_{(n-m) \bmod N} \quad (2)$$

Here,  $L_F$  and  $L_G$  can be defined as two lower triangular matrices which has 1 as the diagonal entry. The first  $(j-1)$  and  $j$  diagonal entries present in  $\tilde{D}$  are compiled to form the diagonal matrices  $D_F$  and  $D_G$ . Determinants of matrices  $F$  and  $G$  can be given by Eqs. 3 and 4 [15],

$$\det\{\mathbf{F}\} = \det\{L_F\} \det\{D_F\} \det\{L_F^H\} = \tilde{D}_1 \tilde{D}_2 \dots \tilde{D}_{j-1} \quad (3)$$

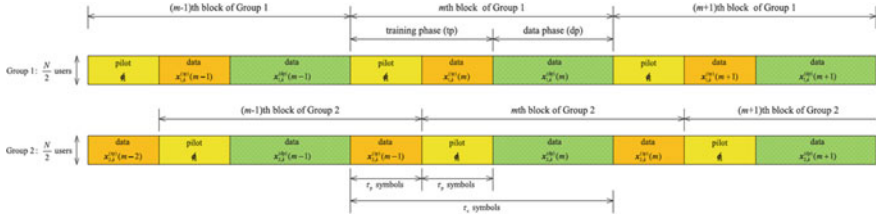
$$\det\{\mathbf{G}\} = \det\{L_G\} \det\{D_G\} \det\{L_G^H\} = \tilde{D}_1 \tilde{D}_2 \dots \tilde{D}_j \quad (4)$$

Using a statistical prototype with optimal equalization coefficients that performs channel estimation, parallel and successive interference cancellation, total of the mean square errors (sum-MSE) of multiple data streams is being diminished, and robust SC-FDE MIMO schemes are created.

The next study involves performance of MU-MIMO system involving time-offset pilots that use a single-cell under successive interference cancellation and minimum mean-squared error channel estimation. To provide accommodation for multiple users for a specific quantity of pilots, recommended time-offset pilots is used. Hence, users can be split into two divisions and the uplink pilots in one division are transmitted concurrently along with the uplink information with the other division. By applying successive interference cancellation, using correlated interference (correlation between pilots of one division and data from the other division) can be efficiently eradicated. Mathematical results make sure that during a short interval of coherence interval, time-offset pilot method was proven to provide greater performance of implementation than the non-orthogonal pilot method. During training period, the data or pilot symbol was represented in the illustration in Fig. 4 [16].

General representation of received signal at the BS is considered to be a vector of dimensions  $M \times 1$  in a single symbol time can be given as Eq. 5 [16],

$$y = \sum_{k=1}^K (\mathbf{h}_{1,k} \sqrt{p_{1,k}} x_{1,k} + \mathbf{h}_{2,k} \sqrt{p_{2,k}} x_{2,k}) + \mathbf{n} \quad (5)$$



**Fig. 4** Design of time-offset pilot design

where the transmit signal of the  $k$ th user in the  $g$ th group that is standardized to be of unit power, i.e.,  $E\{|x_{g,k}|^2\} = 1$ , is denoted by  $x_{g,k}$  ( $g = 1, 2$ ), whereas  $p_{g,k}$  specifies the actual transmit power.

System ability of uniform circular array (UCA)-based orbital angular momentum is enhanced by employing (OAM)-MIMO configurations with successive interference cancellation [17] because capacity of system in the left-over stream is improved by making use of SIC. Due to the huge power difference in the received power, UCA-based OAM performance enhancement was accomplished.

A two-way full-duplex (FD) MIMO communication system including analog and digital TX-RX beamforming can tackle resource problems in hardware due to analog cancellation [18]. MIMO detection is best portrayed as an example of a quadratic optimization issue. To solve this problem, many techniques are available. However, due to high complexity, a lot of them are not possible to be effectively executed at the moment. A good detection performance is provided by MMSE-SIC algorithm. Therefore, to provide low complexity and solve the MIMO issue, dichotomous coordinate descent (DCD) algorithm is beneficial since it can be used in FPGA architecture. A combined detection algorithm with DCD and SIC is proposed. In comparison with MMSE-SIC algorithm, an improved execution and a much lesser intricacy are observed. BER can be further reduced by DCD-SIC when compared to DCD detector, due to the SIC block. In contrast, comparing both MMSE-SIC detector and DCD-SIC detector, the DCD-SIC detector did not require solving using matrix inversion. This can be a plausible reason as to why the DCD-SIC detector portrays lesser intricacy in comparison with the MMSE-SIC algorithm. This increases the chance of making the DCD-SIC algorithm appropriate for implementation using FPGA hardware in huge MIMO systems [19].

Polarization modulation (PM) utilized in massive MIMO-OFDM uplink is being introduced, and exploitation of zero-forcing detection takes place. Moreover, on each sub-carrier, the instantaneous signal-to-noise ratio (SNR) is derived, and examination of the ergodic capacity takes place. It has also been observed in the large PN scenario that PMPNC scheme provides far better performance in comparison with TU scheme, while the number of base station antennas is high. The simulation results [20] show that a higher ergodic capacity can be achieved and the phase noise can be efficiently mitigated through the proposed scheme. Thus, interference cancellation in different types of MIMO systems have been analyzed and surveyed.

## 4 Conclusions

In this survey, for spectral coexistence, the transmission architectures of the multi-input (MIMO) and MIMO radar systems are considered. The initial design is intended to solve two convex optimization problems. An alternative iterative method algorithm will be developed, the solutions of which are given in a closed form that indicate an acceptable trade-off between maximizing the rate of contact and regulating other desired properties of radar waveforms. As a future mission, if they share the same spectrum band, researchers of both the MIMO radar and communication can discuss the effect of CSI errors along with the construction of a stable version of the proposed structure. Furthermore, new and improved techniques for successive interference cancellation in MIMO based wireless body network has been proposed and simulated. Suitable organization of antenna and connections to balance the system, architecture for maximizing complexity reduction, a pre-filter to improve synchronization that helps in enhancing the effect of perfect interference cancellation has been discussed. However, there are some complexity issues and limitations to this study when they are being implemented on hardware-based systems. This study is completely theoretical design based and has to be incorporated in MIMO hardware systems for beneficial results.

## References

1. Saxena A, Khadagade S, Sahu A (2015) A literature review for performance analysis of MIMO-OFDM System. *Int J Innov Trends Eng (IJITE)* 11(01)
2. Zhang J, Glazunov AA, Yang J, Chu X, Zhang J (2018) An experimental study on indoor massive 3D-MIMO channel at 30-40 GHz band. In: 2018 international symposium on antennas and propagation (ISAP). Busan, Korea (South), pp 1–2
3. Chen Y-H, Wong K-L, Li W-Y (2019) 4 X 4 MIMO performance of two conjoined dual wideband antennas including the feedline effects for 5G smartphones. In: 2019 IEEE Asia-Pacific microwave conference (APMC). Singapore, Singapore, pp 1488–1490
4. Chen C, Zhong W, Yang H, Du P (2018) On the performance of MIMO-NOMA-based visible light communication systems. *IEEE Photonics Technol Lett* 30(4):307–310
5. Liu X, Huang T, Shlezinger N, Liu Y, Zhou J, Eldar YC (2020) Joint transmit beamforming for multiuser MIMO communications and MIMO radar. *IEEE Trans Signal Process* 68:3929–3944
6. Hussein HS, Hussein S, Mohamed EM (2020) Efficient channel estimation techniques for MIMO systems with 1-bit ADC. *China Commun* 17(5):50–64
7. Kalachikov AA, Shelkunov NS (2018) Performance evaluation of the detection algorithms for MIMO spatial multiplexing based on analytical wireless MIMO channel models. In: 2018 XIV international scientific-technical conference on actual problems of electronics instrument engineering (APEIE). Novosibirsk, pp 180–183
8. Biswas S, Singh K, Taghizadeh O, Ratnarajah T (2018) Coexistence of MIMO radar and FD MIMO cellular systems With QoS considerations. *IEEE Trans Wireless Commun* 17(11):7281–7294
9. Zhang S, Jin S, Wen C, He Z (2018) Improving expectation propagation with lattice reduction for massive MIMO detection. *China Commun* 15(12):49–54

10. Rihan M, Huang L (2018) Optimum co-design of spectrum sharing between MIMO radar and MIMO communication systems: an interference alignment approach. *IEEE Trans Veh Technol* 67(12):11667–11680
11. Qian J, He Z, Huang N, Li B (2018) Transmit designs for spectral coexistence of MIMO radar and MIMO communication systems. *IEEE Trans Circ Syst II Express Briefs* 65(12):2072–2076
12. Knox ME (2019) Passive interference cancellation in a  $2 \times 2$  STAR MIMO antenna network. In: 2019 International workshop on antenna technology (iWAT). Miami, FL, USA, pp 221–224
13. Kong BY, Jung J, Park I (2018) Interference cancellation architecture for pipelined parallel MIMO detectors. In: 2018 25th IEEE international conference on electronics, circuits and systems (ICECS), Bordeaux, pp 77–80. <https://doi.org/10.1109/ICECS.2018.8618036>
14. Ni Y, Wang Y, Zhu H (2019) Interference cancellation in device-to-device communications underlying MU-MIMO cellular networks. *China Commun* 16(4):75–88
15. Zhe P, Zhu Y, Letaief KB (2018) Robust single-carrier frequency-domain equalization for broadband MIMO systems with imperfect channel estimation. *IEEE Trans Wireless Commun* 17(7):4432–4446. <https://doi.org/10.1109/TWC.2018.2825340>
16. Nguyen TK, Nguyen HH, Nguyen TH (2019) Multiuser massive MIMO systems with time-offset pilots and successive interference cancellation. *IEEE Access* 7:132748–132762. <https://doi.org/10.1109/ACCESS.2019.2940458>
17. Saito S, Suganuma H, Ogawa K, Maehara F (2019) Performance enhancement of OAM-MIMO using successive interference cancellation. In: 2019 IEEE 89th vehicular technology conference (VTC2019-Spring). Kuala Lumpur, Malaysia, pp 1–5. <https://doi.org/10.1109/VTCSPRING.2019.8746501>
18. Iimori H, Thadeu Freitas de Abreu G (2018) Two-way full-duplex MIMO with hybrid TX-RX MSE minimization and interference cancellation. In: 2018 IEEE 19th international workshop on signal processing advances in wireless communications (SPAWC). Kalamata, pp 1–5. <https://doi.org/10.1109/SPAWC.2018.8445776>
19. Tian T, Liu J, Quan Z (2017) Dichotomous coordinate descent based successive interference cancellation algorithm for MIMO systems. In: 2017 IEEE 9th international conference on communication software and networks (ICCSN). Guangzhou, pp 478–481. <https://doi.org/10.1109/ICCSN.2017.8230158>
20. Nie Y, Feng C, Liu F, Guo C, Wu H (2018) Polarization modulation based phase noise cancellation for massive MIMO-OFDM systems. In: 2018 IEEE 29th annual international symposium on personal, indoor and mobile radio communications (PIMRC). Bologna, pp 1–7. <https://doi.org/10.1109/PIMRC.2018.8580988>

# Design and Study the Performance of Micro-Strip Patch Antennae for 5G Mobile Communication



Shivani Krishna Kompella and A. R. Abdul Rajak

**Abstract** The need to update antenna designs to make them more effective has exponentially grown owing to the speedy growth of wireless and radio telecommunication systems. The simulation and analysis of multiple antennae (rectangular, circular, Vivaldi) designed on Rogers RT5880 substrate for future fifth-generation (5G) mobile communications at 40 GHz is reported. The proposed antennae use micro-strip feeding. A modified Vivaldi antenna is designed after running 16 simulations to check for the highest gain output, by changing parameters like size and substrate. The antenna has a maximum gain of 4.763 dB and VSWR of 1.5 at 40 GHz. Because of these improved results, the modified Vivaldi antenna is recommended to employ for 5G mobile communications. Simulation results are obtained using CST Microwave Studio Suite.

**Keywords** 5G communication • Millimeter wave antenna • Line feeding • CST

## 1 Introduction

Due to rapid development in wireless systems and radio telecommunication systems, the need to alter antenna designs to make them more efficient has increased exponentially. Characteristics such as antenna size, bandwidth, power loss, high data rate, and gain have to be improved accordingly in order to achieve a better antenna performance. Fifth generation (5G) technology is almost up to five times more optimal than the usual fourth generation (4G) technology [1]. Hence, it is required to find a perfect trade-off among design, size, cost, and radiation efficiency. The adoption of mm-wave antennae leads to compact sized antennas; however, it suffers from limitations like interference management, anti-blockage, and spatial reuse [2–4].

---

S. K. Kompella (✉)

Birla Institute of Technology and Science, Pilani, Dubai, United Arab Emirates

e-mail: [f20170243@dubai.bits-pilani.ac.in](mailto:f20170243@dubai.bits-pilani.ac.in)

A. R. Abdul Rajak

EEE Department, Birla Institute of Technology and Science, Pilani, Dubai, United Arab Emirates

The frequency spectrum of 5G application is in the range 20–90 GHz, and the antennae designed are usually at frequencies 28, 38, and 72 GHz as they have high data rate, have comparatively low latency, and are highly directional in nature [1]. Due to the narrow beam width designs, these 5G antennae are obstacle sensitive, thereby making them suitable for cellular applications.

Due to their low profile, lightweight, compact scale, low cost, and incredibly basic architecture, micro-strip patch antennae have several widely dispersed functionalities applicable to both non-planar and planar surfaces [5]. The efficiency and gain of the antenna output can be increased by using antenna arrays while designing. Faster access, stronger processors, more power, higher-resolution displays, enhanced Internet connectivity, GPS, third generation (3G), and fourth generation (4G) world access were all the trends for wireless mobile devices, all with longer battery life in smarter, thinner packs.

The 5G mobile communications network offers a far higher efficiency level than previous generations of mobile communications. Next to the next generation of mobile communications, the latest 5G technology progresses from 1 to 2G, 3G, 4G, but it offers a modern solution that gives seamless access. 5G would use ‘huge’ MIMO antenna (multiple inputs, multiple outputs), which have huge numbers of antenna elements or links to send and receive more data simultaneously. The benefit for customers is that more individuals can connect to the network simultaneously and ensure a constant throughput.

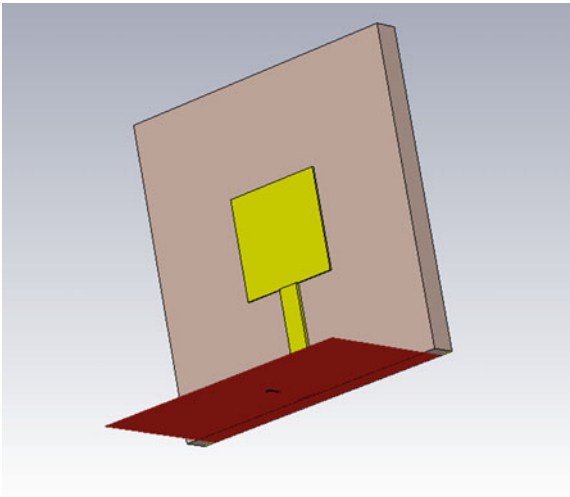
The following antennae are designed: rectangular antenna, circular antenna, and Vivaldi antenna. Comparisons are made based on shape, S-parameters, gain, directivity, and VSWR of the antennae [6].

## 2 Antenna Designs for 5G Communication and Study of Their Performances

### 2.1 Rectangular Antenna Design

A  $L_s \times W_s$  substratum micro-strip patch antenna is with Rogers 5880 material with a 2.2 dielectric constant of 0.0009 loss. Substratum height is 0.508 mm in regular format for Rogers Substrate dimensions are 6 mm  $\times$  18.06 inches. The thickness of the soil is the same as the 6  $\times$  6.25 mm substrate. The feed width is  $W_f$  (0.2 mm) and  $L_f$  (2.15 mm), respectively. The measurements chosen ( $L_p \times W_p$ ) are (2  $\times$  2 mm) for the radiation mask [1, 7] (Fig. 1).

**Fig. 1** Rectangular patch antenna



**Table 1** Dimensions of circular antenna

Component	Dimension	Material
Ground	25 × 25 mm	PEC
Substrate	25 × 25 mm	Rogers RT5880 (lossy)
Circular patch	Outer radius = 6.96 mm	PEC

**2.2 Circular Antenna Design**

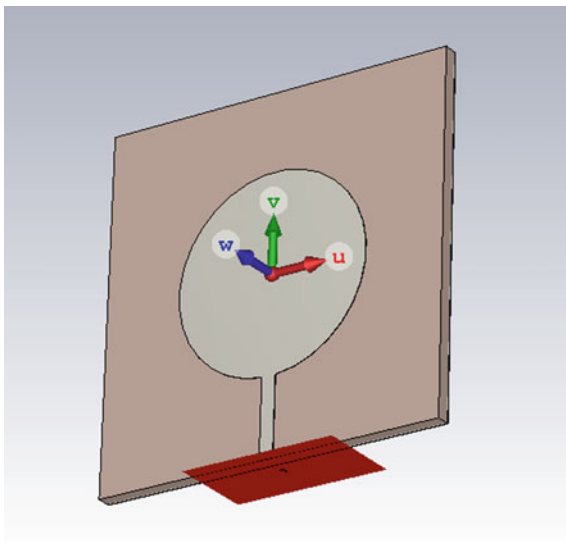
Most of the micro-strip antennas are broadside radiators. The patch is designed in such a way that its pattern is as natural as possible for it. The dimensions of the designed circular antenna are given below in Table 1, along with the diagram as shown in Fig. 2.

**2.3 Vivaldi Antenna Design**

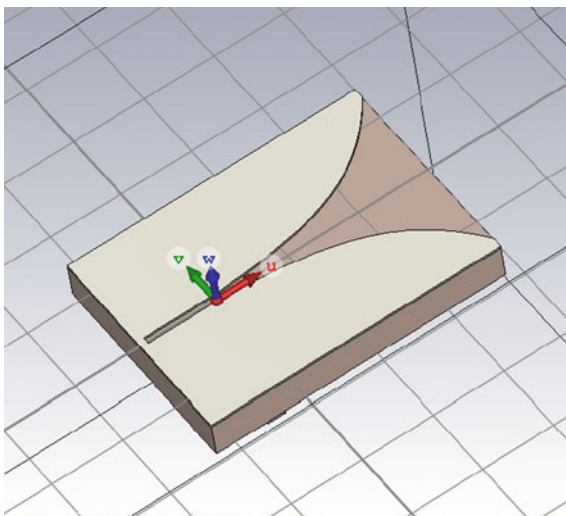
An ethereal or tapered Vivaldi wire or Vivaldi slots antenna is a broadband co-planar receptor which can be manufactured by means of a solid sheet of metal or a printing circuit board or a dielectric plate metallicated on one or both sides. The Vivaldi antennas are helpful for all frequencies, since all reception antennas can be modified to match any recurrence. Printed circuit engineering makes the microwave frequency beyond 1 GHz financially stronger. This antenna (Fig. 3) uses Rogers RT5880 substrate [8].



**Fig. 2** Circular patch antenna



**Fig. 3** Vivaldi antenna design



**Table 2** Gain values for 16 different simulations on changing the parameters of Vivaldi antenna

TL	0.25 (dB)	0.5 (dB)	1.1 (dB)	1.5 (dB)
1	4.227	4.326	4.119	4.053
3	4.763	4.601	3.966	4.412
5	3.504	4.664	4.128	4.685
7	4.225	4.193	4.201	4.621

### 3 Proposal of Modified Structure of Vivaldi Antenna

#### 3.1 Performance of Vivaldi Antenna Parameters Considered for the Modified Structure After Analyzing the Performances

Best substrate: Rogers RT5880

Best feeding: micro-strip line feeding

Frequency: 20–70 GHz (ideally around 30 GHz)

Type: Vivaldi antenna.

#### 3.2 Modified Antenna Design

On simulating the Vivaldi antenna designed previously for multiple values of TL and  $r$ , the following gains are obtained (Table 2).

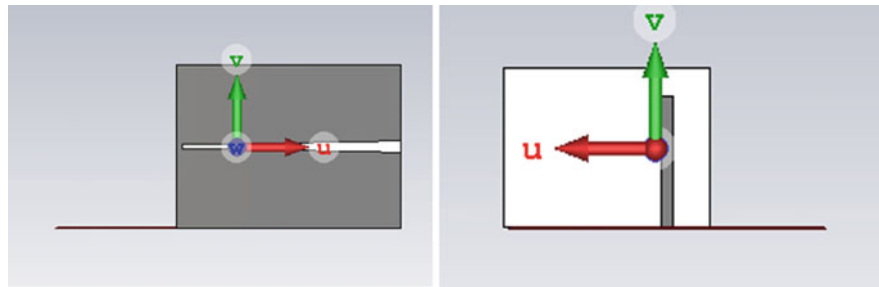
It is seen that increasing TL increases the length of the antenna and increasing  $r$  increases the curve of the substrate layer. Out of these recorded values, at  $TL = 3$  and  $r = 0.25$ , the Vivaldi antenna is seen to exhibit the best gain characteristics. Hence, an updated antenna is designed using these new parameters to obtain optimal results (Figs. 4 and 5).

The new dimensions used for this Vivaldi antenna are mentioned in Table 3.

#### 3.3 Results Obtained on Simulating This Antenna

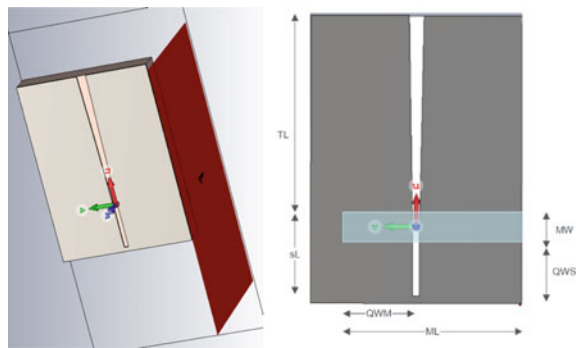
##### 3.3.1 S Parameters

Figure 6 shows that  $S(1,1)$  is  $-13.978$  at 42.72 frequency per GHz central frequency.



**Fig. 4** Front and back view of Vivaldi antenna

**Fig. 5** Vivaldi patch antenna



**Table 3** New updated dimensions of Vivaldi antenna to provide better gain

Name	Value
s	0.05
Ext	0.1
R	0.25
MW	0.25
H	0.5
QWS	0.75
QWM	0.9
sL	1
TL	3
subw	3
k	5.96

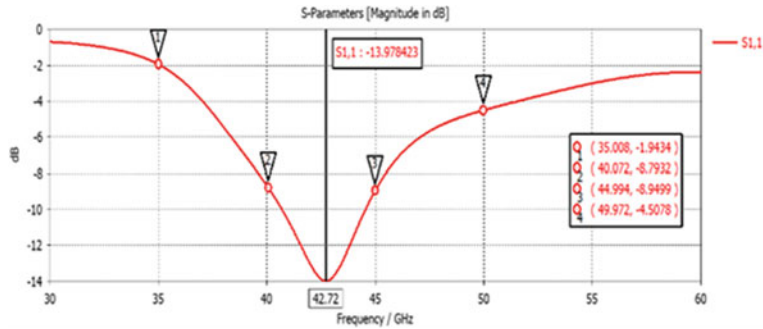


Fig. 6 S-parameters

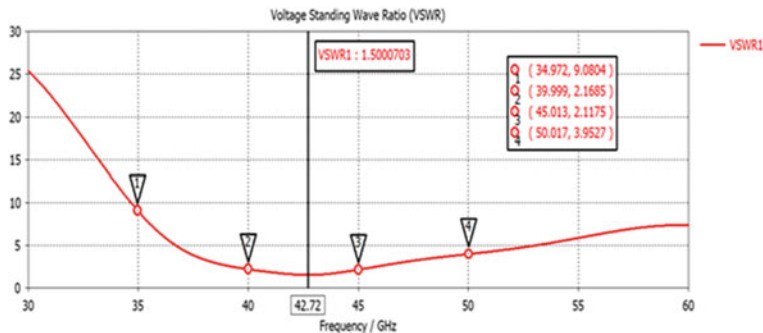


Fig. 7 Voltage standing wave ratio

3.3.2 Voltage Standing Wave Ratio (VSWR)

$VSWR = 1.5$  is obtained, as shown in the graph Fig. 7. According to the study, if the VSWR of any antenna is below 2, the antenna-match is measured to be very good and trying to make it better by impedance matching would not be very rewarding.

As the VSWR keeps increasing, it means that larger power is being reflected from the antenna and not transmitted.

3.3.3 Power Loss and Total Efficiency

Losses in antenna efficiency are normally caused by driving losses due to the finite conductivity of the antenna metal, dielectric losses—as a dielectric material is conductive in the presence of an aerial and mismatch loss impedance. Results are shown in Figs. 8 and 9.

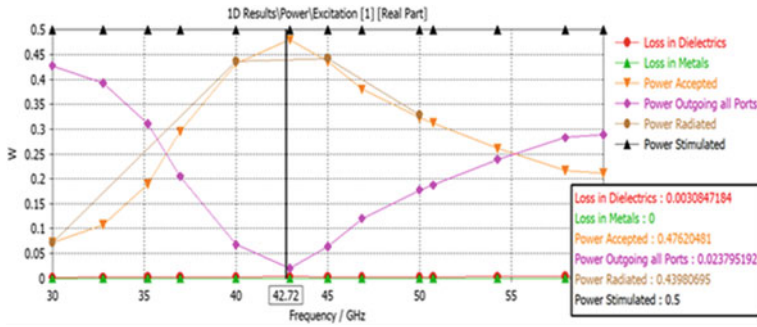


Fig. 8 Power loss

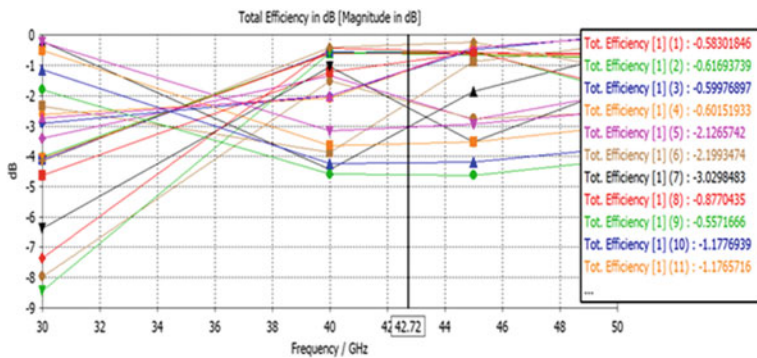


Fig. 9 Total efficiency

### 3.3.4 Antenna Gain

The term antenna gain indicates the power transferred to that of an isotropic source in the direction of peak radiation. From Fig. 10, the Vivaldi antenna gain is seen to be 4.763 dB. An antenna with a 3 dB power gain means that the antenna's power is 3 dB (almost double) higher than what a lossless isotropic antenna would have achieved with the same input power.

## 4 Conclusion

In this research, multiple micro-strip patch antennae are designed for 5G wireless communication. Comparisons were made based on their S-parameters, power loss, efficiency, gain, shape, and substrate. Table 4 summarizes all the key highlights of this study.

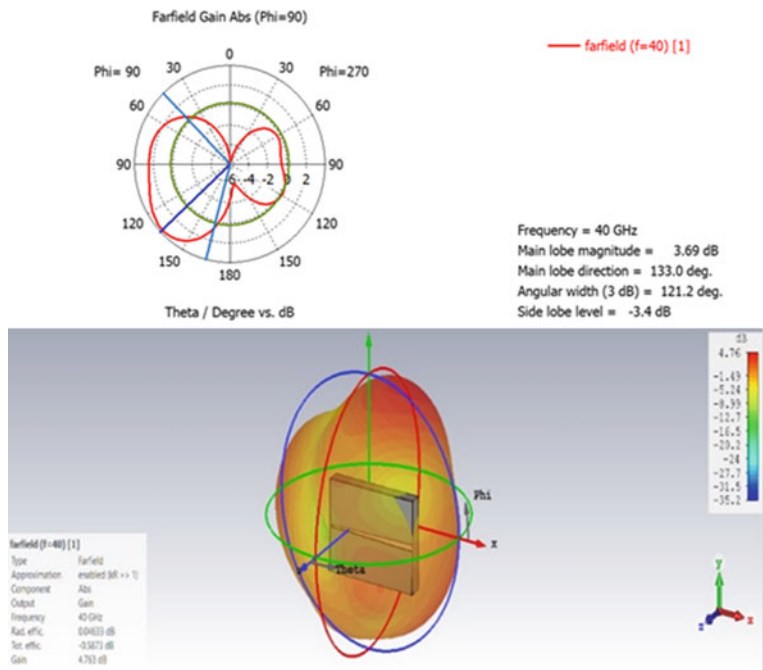


Fig. 10 Far-field (polar form and 3D form)

Table 4 Characteristics of all the antennae designed in this project report

Antennae	Frequency (GHz)	S Parameter (dB)	VSWR	Gain (dB)
Rectangular	40	-7.64	2.418	4.101
Circular	40	-34.33	1	7.387
Vivaldi	40	-14.49	1	3.966
Modified Vivaldi	40	-13.97	1.5	4.763

This shows that high gain can be achieved at higher frequencies by 5G mode of communication. These antennae have higher gain, bandwidth, radiation efficiency, and directivity as made known by the graphical results presented in this report, which illustrates that this millimeter wave micro-strip patch Vivaldi antenna of 40 GHz is very suitable for 5G communications.

## References

1. Imran D, Farooqi MM, Khattak MI, Ullah Z, Khan MI, Khattak MA, Dar H (2018) Millimeter wave microstrip patch antenna for 5G mobile communication. In: 2018 international conference on engineering and emerging technologies (ICEET)
2. Amrutha GM, Sudha T (2019) Millimeter wave doughnut slot MIMO antenna for 5G applications. In: TENCON 2019—2019 IEEE region 10 conference (TENCON), Kochi, India, pp 1220–1224. <https://doi.org/10.1109/TENCON.2019.8929658>
3. El Shorbagy M, Shubair RM, Al Hajri MI, Mallat NK (2017) On the design of millimetre-wave antennas for 5G. In: IEEE microwave symposium (MMS), 2016 16th mediterranean, pp.1–4
4. Al-Ogaili F, Shubair RM (2016) Millimeter-wave mobile communications for 5G: challenges and opportunities. In: Antennas and propagation (APSURSI), 2016 IEEE international symposium, pp 1003–1004
5. Rajput GS Design and analysis of rectangular microstrip patch antenna using metamaterial for better efficiency. *Int J Adv Technol Eng Res (IJATER)*
6. Madhan MD, Subitha D (2019) Millimeter-wave microstrip patch antenna design for 5G. *Int J Innov Technol Explor Eng (IJITEE)* 8(12). ISSN: 2278-3075
7. Jamal A, Parmar R, Razak A, Design and performance analysis of Microstrip Patch Antenna following IEEE 802.11b/g/a specifications. In: 2nd IEEE international conference on human computer interactions (ICHCI'16) held at Nalli Arangam, Saveetha University, Chennai. Tamil Nadu
8. Hasan MN (2018) Dispersion characterization of a UWB Vivaldi antenna in time and frequency domain. In: IEEE Indian conference on antennas and propagation (InCAP)

# Sentiment Analysis of COVID-19 Vaccine Rollout in India



Sushila Paliwal , Suraiya Parveen , M. Afshar Alam ,  
and Jawed Ahmed

**Abstract** This paper deals with opinions expressed in social media, namely Twitter, with respect to the ongoing vaccination program for the novel coronavirus COVID-19. This pandemic has affected almost the entire world because of lockdown and its related impact on personal, economical and emotional levels. During the lockdown, people suffered loneliness and lack of interpersonal interaction therefore resort to taking social media to voice their concerns and emotions. The need for vaccination, its delivery and trust are some of the topics on which sentiments are expressed on Twitter based on which this research is conducted. This study analyses the sentiments of the people in India towards the emergency use authorization of COVID vaccine rollout. The tweets are collected, pre-processed to exclude noise and irrelevant data and then used for sentiment analysis using the Lexicon-based approach and machine learning method. The relative outcome of tweets analysed via all these methods is consistent in each type of sentiment. It is observed that sentiments of most of the tweets came out to be either affirmative or neutral.

**Keywords** COVID-19 · Vaccine · Sentiment analysis · Natural language processing · TextBlob · Vader · Random forest · COVID-19 vaccination

## 1 Introduction

Social platforms such as Facebook, Twitter and Instagram enable people to share their emotions. Twitter has become one of the most popular microblogging platforms to share information, opinion and emotions [1]. Nowadays, users can post (“Tweets”) with a maximum of 280-characters that express thoughts, information, comments and observations, though initially limit was up to 140-characters only [2]. Researchers have used the Twitter platform for analysing the sentiments and emotions in almost

---

S. Paliwal (✉) · S. Parveen · M. Afshar Alam · J. Ahmed  
Department of Computer Science and Engineering, School of Engineering Sciences and  
Technology, Jamia Hamdard, New Delhi, India

M. Afshar Alam  
e-mail: [aalam@jamiahamdard.ac.in](mailto:aalam@jamiahamdard.ac.in)



every sphere of life such as politics, religion, environment, and government policies. This paper discusses the Indian vaccine rollout sentiment analysis of COVID-19.

Globally, 89,416,559 confirmed cases of coronavirus have been reported to the World Health Organization (WHO) of which 1,935,028 were deaths as of 12th January 2021 [3]. India has recorded the third-highest deaths after the USA and Brazil with about 151,327 as of 12th January 2020. A man of 76 years from Kalburgi, Karnataka, with an international travel history, was the first person to die in India from COVID-19 on 12 March 2020 [4]. As there was no prescribed medication to treat infection of coronavirus, vaccine is being developed at a pandemic rate by researchers [5] and all the countries are actively participating in it.

Healthcare decisions are personal and depend on the patients to participate in the specific treatments. In order to plan the vaccine distribution initiatives, it is very important to understand the sentiments of people within specific geographies or demographic communities. Government health agencies and healthcare providers should prepare the distribution strategy based on the general feelings of the population about receiving a vaccination. Sentiment analysis can impact the vaccination initiative and changes in beliefs and behaviour of people towards the vaccinations [6].

Vaccination-related sentiments can affect the individual decisions of taking the vaccine. Measuring these sentiments is a resource-intensive and difficult task. We use data from Twitter to measure the sentiments towards the emergency use authorization of two vaccines in the early phase of vaccination rollout. In India, “Emergency Use Authorization or Accelerated Approval has been granted by the National Regulator for two vaccines—Covishield and Covaxin—which have established safety and immunogenicity on 3rd January 2021” [7]. The vaccination drive of COVID-19 in India is planned to begin on 16th January with an estimated three crore healthcare workers and frontline employees prioritized [8]. It added that this would be followed by voluntary vaccinations of those over the age of 50 and of co-morbidity population groups under 50, comprising around 27 crores.

A great resource for learning about people’s feelings is the widespread emotion on the pandemic vaccination and analysing sentiment of vaccination over time will keep your finger on the pulse of people’s thoughts and feelings. Sentiment analysis [9] is an integral part of the monitoring of social media because it enables us to take a look at the prevailing opinions behind a given subject. It is true that humans can interpret behaviour of people much better than a machine but that is not possible with a large amount of data; however, it can be achieved by combining natural language processing and machine learning.

## 2 Related Work

Natural calamities and pandemic-related Internet community research have been done using various natural language processing (NLP) techniques. Sentiment analysis is one of the most investigated NLP topics among these different techniques.

Many researchers have analysed public sentiments based on social media data towards vaccination. During the outbreak of novel influenza A (H1N1) [10, 11], collected publicly available tweets for six months and a classifier has been created to predict sentiments for tweets automatically. They established that “projected vaccination rates based on sentiments expressed on Twitter are in agreement with vaccination rates estimated by the CDC with traditional phone surveys”. A study conducted by Gloria Kang et al. [12] showed a terminology that was used in framing positive and negative messages within vaccine communication. Public opinion about HPV vaccines is analysed by Du et al. [13] using the sentiment analysis based on machine learning. They worked on sentiment analysis for newly approved HPV vaccines using Twitter data to extract public opinion automatically and instantly to understand their concerns.

Researchers from all over the world are analysing Twitter data to find people’s reactions to the issues such as lockdown, safety measures due to Corona. Sentiments and emotions of the people of twelve countries during the pandemic COVID-19 are analysed to know how individuals in those countries are reacting to the outbreaks of the disease [14]. The results of the study conclude that majority of the people are positive and hopeful but there are instances of fear, sadness and distress revealed worldwide. The citizens of India have positively welcomed the initial lockdown of 21 days declared by the Government of India [15]. Top words in the tweet dataset reflect public emotions about the lack of food and fear associated with it. Sentiment and emotion of peoples in the USA investigated on the subject of reopening [16]. In this work, they try to get insights into public reaction as the reopen phase starts and found that even though people are showing a positive attitude to reopen their areas to make the economy functional. People are also requesting the authority in concern to make plans to avoid the highly predictable second wave.

### 3 Dataset

The tweets have been collected using the Python Library Tweepy [17] from Twitter. It allows appropriate data retrieval by searching via keywords, hashtags, trends, or geo-location. We collected available Twitter text messages (called “tweets”) with at least one of these search strings: “Covishield”, “Covaxin”, “Vaccine” or “Immunization”. For this research, keywords in the English language have been used to extract tweets rather than hashtags as keyword approaches are more inclusive because they also contain the hashtag tweets [18]. We extracted the tweet text, time and date of the publication of the tweet and user location (if available). These tweets have been gathered from 10th January 2021 to 16th January 2021 on the day of vaccine rollout in India. A total of 12,482 tweets have been collected from Twitter for Indian locations, pre-processed and then used for sentiment analysis. Data extracted from Twitter might have duplicates because of same tweets are collected again by API or when multiple users post/retweet the same tweet. After dropping duplicates rows and unnecessary columns, 7640 tweets were used for our analysis.

## 4 Methodology for Sentiment Analysis

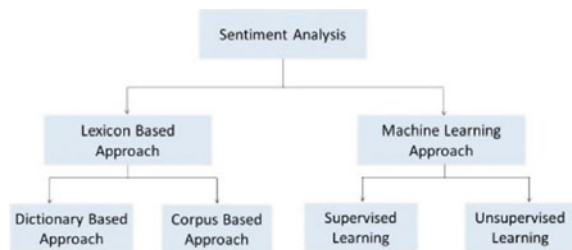
Sentiment classification techniques are divided into lexicon-based approach [19] and machine learning approach [20] as shown in Fig. 1. The lexicon-based approach is further divided into dictionary-based and corpus-based approaches. The lexicon approach is based on a lexicon of sentiment, which is a collection of known and precompiled terms of sentiment.

The dictionary-based approach uses an online dictionary to tag words, and the corpus-based approach depends on statistics of co-occurrence or syntactic patterns embedded in text corpora. A linguistic resource consisting a priori information about words is the final outcome throughout the semantic dimension of sentiment. A sentiment lexicon is a dictionary of positive and negative words, positive words could be “joy”, “great”, etc., while negative words could be “terrible”, “cruel”, etc. The occurrence of positive and negative words in the text is counted, and a positive or negative score is assigned to it based on the counts using statistical or semantic methods for finding sentiment polarity. The lexicon-based approach is easier to implement as compared to the machine learning-based approach as it does not require labelling the data for training. However, the machine learning-based approach has the benefit of training the models with domain-specific data to analyse a given text which increases the accuracy and confidence level in sentiment analysis [21].

### 4.1 *TextBlob and Vader Are Dictionary-Based Sentiment Analysis Tools Which We Have Used in This Research*

- TextBlob is a Python Library built upon NLTK and is used for processing textual data [22]. This library offers a simple API to access common natural language processing (NLP) tasks. NLP [23] tasks include part-of-speech tagging, sentiment analysis, noun phrase extraction, classification, translation, etc. TextBlob is just like a Python string hence very easy to use. TextBlob’s sentimental function returns polarity and subjectivity. Polarity is a float value within the range  $[-1.0, 1.0]$ , in which  $-1.0$  is a negative statement and  $1.0$  a positive statement. Subjective phrases usually refer to personal opinions, emotions or judgements, whereas objective

**Fig. 1** Sentiment analysis methods



refers to facts. Subjectivity property is also a float value within the range of [0.0, 1.0] where 0.0 is very objective and 1.0 is very subjective.

- Valence Aware Dictionary and sEntiment Reasoner (VADER) is a lexicon and rule-based sentiment analyses tool [24]. To specify a mathematical model without explicitly encoding VADER uses a set of rules and is a model of sentimental analysis with less resource consumption. The study shows that VADER works equivalently to individual human raters for matching ground truth as it is specifically tailored for sentiments on social media [25]. VADER analyses part of the text in order to verify that any word in the text is available in the VADER lexicon. Polarity indices are found using the `polarity_scores()` function. This returns metric values for a particular sentence: negative, neutral, positive and compound. The compound score is a method that computes the total of all lexicon ratings normalized from  $-1$  to  $+1$  in which  $-1$  shows the most extreme negative ratings and  $+1$  shows the most extreme positives. It is useful to set standard thresholds as positive, neutral or negative for the classification of sentences. The following are the typical thresholds:
  - Positive Sentiment: compound score  $\geq 0.05$
  - Negative Sentiment: compound score  $\leq -0.05$
  - Neutral Sentiment: compound score  $> -0.05$  and  $< 0.05$

#### ***4.2 Random Forest is a Machine Learning-Based Approach Which We Have Used in This Research***

- Different machine learning classifiers have their own set of parameters which are required to be taught before the model gets trained to get the best results. Random forest uses multiple decision trees and hence reduces the risk of overfit. It is good for large datasets and can be trained relatively quickly, can be used for both regression and classification problems and works fine with missing data. Random forest is an ensemble learning method [26] that works by creating several decision trees during the training stage. For the classification problem, the random forest makes a final decision based upon the decisions chosen by the majority of the trees. In the training phase, random forest built multiple decision trees that are used to build a single model, and in the testing phase, we used these models to predict the outcome.

## **5 Experimental Evaluation**

This section explains the data pre-processing and experimental setup of the sentiment analysis.

## 5.1 *Twitter Connection*

In order to access the tweets, we created a Twitter-developer account and after authentication, we extracted the tweets by using `API.search()` method to fetch the tweets by a keyword.

## 5.2 *Collection of Tweets*

Collected the tweets through Twitter API using Tweepy library of Python with keywords “Covishield”, “Covaxin”, “Vaccine” or “Immunization”.

## 5.3 *Pre-processing of Tweets*

Usually, data received from Twitter is unstructured and contains unnecessary symbols which are not essential for our analysis. Pre-processing is the mandatory step for cleaning the raw tweets by removing noise and unwanted attributes before conducting sentiment analysis. We have used stop word removal, tokenization and text normalization to pre-process the tweets.

Stop word removal—Common words that occur in sentences that add weight to the sentence are known as stop words such as “of”, “the” and “for”. These stop words act as a link and make sure that sentences are grammatically correct. Filtered out words before processing natural language data are known as stop words, and it is a common pre-processing method.

The tokenization method is used to divide a sentence, a phrase, a paragraph or whole text into smaller parts. By doing that, we can get the individual terms or words and each of these individual terms is known as tokens. It includes:

- Remove irrelevant words: Removing non-ASCII characters and irrelevant words (i.e. RT, &amp) from the tweet. All these symbols are replaced with a whitespace character.
- Remove mentions and hashtags: Cleaning the data of mentions and hashtags will help in reducing the noise.
- Remove URL and Email: Remove URLs and emails using regular expression.
- Remove duplicates which do not add any value.
- All the texts in tweets are converted into lower case.

Text normalization—Text must be normalized before further processing as normalization improves text matching. Normalization generally refers to that words are allowed on an equal footing and allow the processing to continue uniformly. Stemming and lemmatization [27] are the two tasks used to normalize the text:

- **Stemming:** Eliminating affixes from a word for a word stem is known as stemming. Porter stemmer is very fast; hence, it is the most widely used technique. Typically, the end of the word chops off in stemming and it works well most of the time. For example, words like vaccinate, vaccinated, vaccination and vaccinating were stemmed to the word “vaccine”.
- **Lemmatization:** The intent is the same as stemming but occasionally loses the meaning of the word through stemming. To do things properly with vocabulary and morphological word analysis is called lemmatization. The base or a dictionary form of a word is returned by it which is known as lemma. Lemma is an actual language word returned by lemmatization while stem might not be an actual word returned by stemming. For instance, words like coronavirus, ncov and virus were transformed into its citation form to “COVID”. It reduces the distinct number of words in the text and will improve the analysis.
- Stemming and lemmatization both are normalization techniques, and we have used the lemmatization approach to normalize as Lemmatization is a proper approach for doing things with a vocabulary and a morphological word analysis.

## ***5.4 Extracting Features from Cleaned Tweets***

The method for extracting out a list of words from textual data is known as text extraction feature [28]. The list of word features needs to be extracted from the tweets and transformed into a feature set. Before we implement ML text classifier, we need the relevant features set.

## ***5.5 Modelling Technique***

The process of converting tokens to numbers is known as vectorization, and it is an essential step as the algorithm expects data in some mathematical form rather than textual form. Many techniques such as N-grams, TF-IDF and bag of words are available for vectorization. We have implemented vectorization using scikit’s TF-IDF vectorizer [29]. TF-IDF is a numerical statistical method to determine how a word is important to a corpus, and it is also used to reflect how relevant a term is in a given document. This process compares the frequency of the word in the document against the frequency in the entire document group (corpus).

## ***5.6 Model Building***

We are now done with all the pre-modelling stages required to get the data in the proper form and shape. We will be building a model on the datasets with the TF-IDF




most prominent words as: “Vaccine”, “Covaxin”, “India”, “Covishield” and “COVID vaccine” which represents the sentiments of Indian people for vaccination.

In this research, the data of Twitter was retrieved between 10th Jan and 16th Jan 2021 by Twitter API (Tweepy) and was pre-processed for any ambiguities. Tweets were analysed using three standard computational linguistic models—two are based on lexicon (TextBlob and Vader) and the third one is a random forest algorithm-based machine learning model to process the cleaned data. Example tweets with each sentiment value are shown in Fig. 3.

While doing sentiment analysis using VADER, the graph of day-wise sentiments during the study period is shown in Fig. 4. It shows that on daily basis negative sentiments are consistently less as compared to positive and neutral sentiments. Here we can also see that negative sentiments have been the least as compared with positive and neutral sentiments. The graph of TextBlob analysed sentiments in Fig. 5 illustrates that the positive sentiments are slightly higher in comparison with the neutral sentiments during the study period.

The random forest algorithm consists of several separate independent decision trees which have been independently trained based on a random subset of data.

Fig. 3 Example tweets

Sentiments	Example Tweets
	Proud day for we Indians 🇮🇳. Our scientists have been working tirelessly to roll out indigenous #VaccineForIndia. Finally successful in rolling out Covaxin and Covishield.
	It took 10 years to produce a vaccine and bring it to the market before 2019. Now, it is less than 1. Something is wrong. #vaccination #vaccine #vaccinationCovid #CovidVaccine #COVID19
	COVID-19 vaccination to begin this week: Everything you need to know

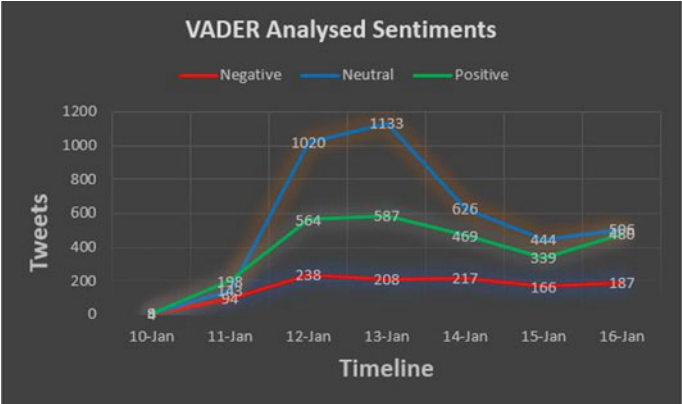


Fig. 4 VADER analysed sentiments



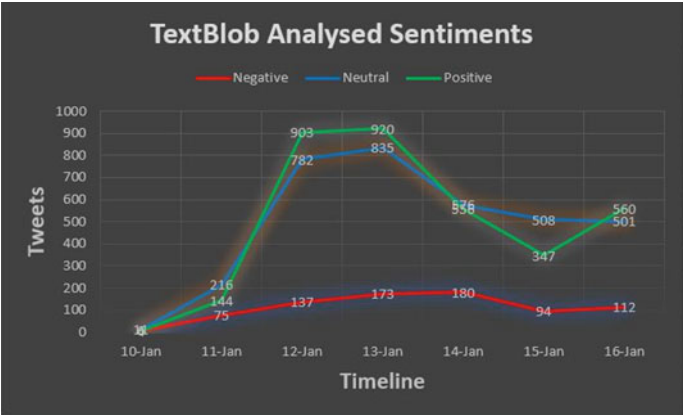


Fig. 5 TextBlob analysed sentiments

Random forest day-wise representation of the data is shown in Fig. 6. From the graph of random forest analysed sentiments, we can conclude that the sentiments are in line with the results of VADER and TextBlob. Overall negative sentiments are the lowest, and neutral sentiments are the highest

The comparative analysis of all the three approaches on the same dataset is shown in Table 1. We used a random forest classifier to feed some of the Tweets that are already manually classified then we used it to find the sentiment value for new Tweets.

Sentiments analysed by TextBlob, Vader and random forest are shown percentage-wise in Fig. 7. The analysis of VADER showed 34.6% positive tweets, 14.6% negative

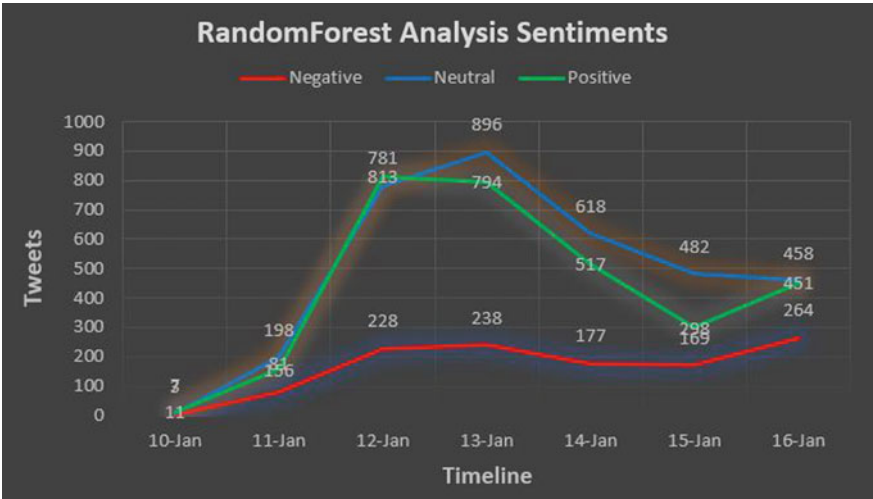
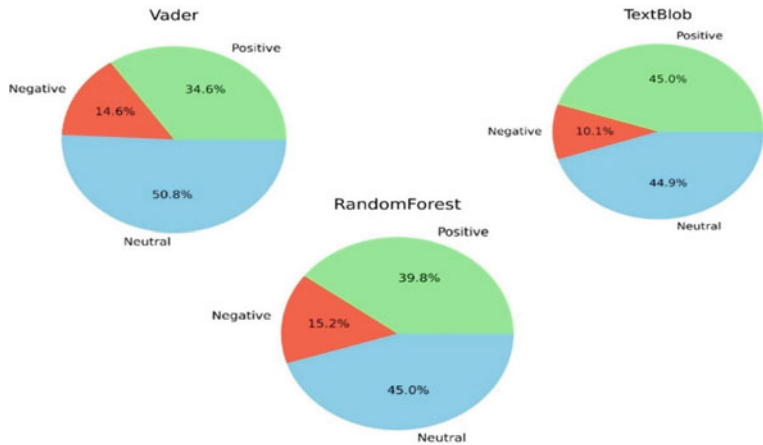


Fig. 6 Random forest analysed sentiments

**Table 1** Sentiment values by three methods

Sentiments	Vader	TextBlob	Random forest
Neutral	3880	3429	3440
Negative	1119	775	1160
Positive	2641	3436	3040



**Fig. 7** Overall sentiments for vaccination during the study period

tweets and 50.8% neutral tweets, as for the TextBlob analysis, the tweets were 45.0% positive, 10.1% negative and 44.9% neutral. Furthermore, random forest analysed 39.8% positive, 15.2% negative and 45.0% neutral tweets. It can be seen that the majority of tweets have a neutral sentiment followed by positive sentiments, and overall, there are fewer negative sentiments related to vaccination.

7 Conclusion

Social media is heavily being used for information and advice; hence, mining of this data is trending research these days. Data available on social media is too huge to analyse manually. There are known methods for analysing this data using techniques based on lexicon and based on machine learning. In this work, we have interpreted public reactions on vaccine rollout for essential services and frontline workers by using Vader, TextBlob and random forest. Vaccine rollout comments/opinion of the public was too vague to analyse so in the methodology we have used pre-processing of the rollout data available on Twitter using tokenization, stop word removal and text normalization (stemming and lemmatization) technique of NLP. The analysis gives a clear-cut indication of the public opinion on COVID vaccination in India that there is no negative wave for the vaccination but the response from the public

is not so overwhelming. A vaccine of COVID-19 was the most awaited moment, when the entire world was in the grip of the Corona pandemic, all voices in the social networking sites (in its all form) were in dire need of vaccine. But when finally, the vaccine was unveiled and made available to frontline workers/worriers, its demand is diminishing, and mass public and frontline workers have not shown much enthusiasm. We got this conclusion using the lexicon-based technique and the random forest algorithm of machine learning. We found that all these three techniques are giving us the same kind of results. Sentiment analysis for the newly approved corona vaccine identifies the concerns and provides an insight into the viewpoint of the public by using Tweets. This analysis can provide feedback for health professionals to effectively plan their vaccine promotion strategies.

## References

1. von Nordheim G, Boczek K, Koppers L (2018) Sourcing the sources: an analysis of the use of Twitter and Facebook as a journalistic source over 10 years in The New York Times, The Guardian, and Süddeutsche Zeitung. *Dig J* 6(7):807–828. <https://doi.org/10.1080/21670811.2018.1490658>
2. Tweet length. <https://developer.twitter.com/en/docs/counting-characters>. Accessed 14 Feb 2021
3. World Health Organization. <https://covid19.who.int/>. Accessed 14 Feb 2021
4. India's first coronavirus death is confirmed in Karnataka. <https://www.hindustantimes.com/india-news/india-s-first-coronavirus-death-in-karnataka-confirmed/story-2ZJ6IuxJ38EiGndBq5pfHO.html>. Accessed 14 Feb 2021
5. Lurie N, Saville M, Hatchett R, Halton J (2020) Developing Covid-19 vaccines at pandemic speed. *N Engl J Med* 382(21):1969–1973. <https://doi.org/10.1056/nejmp2005630>
6. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J (2012) Misinformation and Its correction: continued influence and successful debiasing. *Psychol Sci Public Interes Suppl* 13(3):106–131. <https://doi.org/10.1177/1529100612451018>
7. Press Information Bureau (PIB). <https://pib.gov.in/PressReleasePage.aspx?PRID=1685763>. Accessed 14 Feb 2021
8. Press Information Bureau (PIB). <https://pib.gov.in/Pressreleaseshare.aspx?PRID=1687305>. Accessed 14 Feb 2021
9. Joshi A, Bhattacharyya P, Ahire S (2017) Sentiment resources: lexicons and datasets
10. 2009 swine flu pandemic. [https://en.wikipedia.org/wiki/2009\\_swine\\_flu\\_pandemic](https://en.wikipedia.org/wiki/2009_swine_flu_pandemic). Accessed 14 Feb 2021
11. Salathé M, Khandelwal S (2011) Assessing vaccination sentiments with online social media: Implications for infectious disease dynamics and control. *PLoS Comput Biol* 7(10). <https://doi.org/10.1371/journal.pcbi.1002199>
12. SS, Gloria Kang GJ, Ewing-Nelson SR, Mackey L, Schlitt JT, Marathe A, Abbas KM (2018) 乳鼠心肌提取 HHS Public Access. *Physiol. Behav* 176(1):139–148. <https://doi.org/10.1016/j.vaccine.2017.05.052.Semantic>.
13. Du J, Xu J, Song HY, Tao C (2017) Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. *BMC Med Inform Decis Mak* 17(Suppl 2). <https://doi.org/10.1186/s12911-017-0469-6>
14. Akash DD (2020) Twitter sentiment analysis during Covid-19 outbreak. *SSRN Electron J March*:1–9. <https://doi.org/10.2139/ssrn.3624719>
15. Kaur R, Ranjan S (2020) Sentiment analysis of 21 days COVID-19 Indian lockdown tweets. *Int J Adv Res Sci Eng* 9(7):37–44

16. Ahmed ME, Rabin MRI, Chowdhury FN (2020) COVID-19: social media sentiment analysis on reopening. arXiv
17. Tweepy Documentation. <https://docs.tweepy.org/en/latest/>. Accessed 14 Feb 2021
18. Pulido CM, Villarejo-Carballido B, Redondo-Sama G, Gómez A (2020) COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information. *Int Sociol* 35(4):377–392. <https://doi.org/10.1177/0268580920914755>
19. Taboada M, Brooke J, Voll K (2011) Lexicon-based methods for sentiment analysis
20. Gopi AP, Jyothi RNS, Narayana VL, Sandeep KS (2020) Classification of tweets data based on polarity using improved RBF kernel of SVM. *Int J Inf Technol*. <https://doi.org/10.1007/s41870-019-00409-4>
21. Gbashi S, Adebo OA, Doorsamy W, Njobeh PB (2021) Systematic delineation of media polarity on COVID-19 vaccines in Africa: computational linguistic modeling study. *JMIR Med Inf* 9(3):e22916. <https://doi.org/10.2196/22916>
22. Loria S (2020) TextBlob documentation. TextBlob, p 69, [Online]. Available <https://textblob.readthedocs.io/en/dev/>
23. Kang Y, Cai Z, Tan CW, Huang Q, Liu H (2020) Natural language processing (NLP) in management research: a literature review. *J Manag Anal* 7(2):139–172. <https://doi.org/10.1080/23270012.2020.1756939>
24. Hutto CJ, Gilbert EE (2014) VADER: a parsimonious rule-based model for sentiment analysis of social media text. eighth international conference on weblogs and social media (ICWSM-14). In: Proceeding 8th International Conference Weblogs Social Media, ICWSM 2014. [Online]. Available: <http://sentic.net/>
25. Elbagir S, Yang J (2019) Twitter sentiment analysis using natural language toolkit and Vader sentiment. *Lect Notes Eng Comput Sci* 2239:12–16
26. Pal R (2017) Overview of predictive modeling based on genomic characterizations. *Predict Model Drug Sensit* 121–148. <https://doi.org/10.1016/b978-0-12-805274-7.00006-3>
27. Anjali GJ, Anjali M (2007) A comparative study of stemming algorithms. *October* 2(2004):1930–1938
28. Waykole RN, Thakare AD (2018) A review of feature extraction methods for text classification. *Int J Adv Eng Res Dev* 5(04):351–354
29. Li H, Phung D (2014) Scikit-learn machine learning in Python. *J Mach Learn Res* 39(2014):i–ii
30. WordCloud for Python documentation. [https://amueller.github.io/word\\_cloud/](https://amueller.github.io/word_cloud/). Accessed 14 Feb 2021

# Detection of Swine Flu System Using IOT and Machine Learning



Devyani Bhamare and V. V. Mandhare

**Abstract** Pig influenza is an arising viral disease that addresses an ebb and flow worldwide general medical issue. There are a huge number of instances of pig influenza today. This new disease can be seen everywhere on the present reality. This disease is a sort of variation of the H1N1 influenza contamination. In India, consistently the table of tainted individuals is expanding, so it is imperative to accept this illness into account as it very well may be deadly. Wellbeing frameworks engineering for machine learning and IOT is arising as a significant objective with the prosperity of better caliber of administration, cost decrease and adaptability. As we probably are aware, AI holds new guarantee for clinical analytic applications. Consequently, this article presents the idea of pig influenza finding fixated on counterfeit neural organizations (ANNs). ANN characterizes a default rating to separate certain circumstances and the framework isolates positive or negative experiments dependent on the standard framework. The proposed model uses an accessibility calculation to choose the fundamental attributes for the pig influenza indicative test utilizing the arbitrary backwoods classifier and the C5 calculation. This is on the grounds that it diminishes the length of the data that will be utilized to prepare the ANN model with a substance.

**Keywords** IOT · Machine learning · ANN

## 1 Introduction

Pig influenza is quite possibly the most irresistible illnesses that contains kinds of infections that cause a huge number of passing each year. Pig influenza is a respiratory illness brought about by flu infections that taint the respiratory lot of pigs and cause a yapping hack, diminished hunger, runny nose and languid conduct; the infection can be sent to people. Pig seasonal infections can transform or change so

---

D. Bhamare (✉) · V. V. Mandhare  
Sanjivani College of Engineering, Kopergaon, India

V. V. Mandhare  
Pravara Rural Engineering College, Loni, India

they are handily sent between people. Pig influenza, likewise called pig influenza or pig influenza, pig influenza and pig influenza, is a disease brought about by any of the various sorts of pig seasonal infection. Some regular side effects of pig influenza are: hack, fever, sore throat, runny nose, cerebral pain, chills, weariness, queasiness. For a more prominent improvement and exactness in the outcome, we utilized the IOT area framework since it has been progressively powerful in wellbeing programs. Across the medical care region, the IOT generally consolidates gadgets like microcontrollers to decipher and communicate sensor information to the worker and furthermore to doctors. The coordination of this IOT innovation into clinical gear builds the effectiveness and conveyance of treatment for both the old and youngsters. Consequently, in the medical services area, the IOT can store a huge number of patient information and permit admittance to their information whenever. Then again, there have been wellbeing screening gadgets that are minimized and permit patients to utilize them for observing. The wellbeing observing framework is associated with the patient so the specialist can screen the patient's status whenever. Since IOT-minded patients can be reached on the Internet, it is essential to screen the patient's wellbeing status at the correct time so fundamental moves can be made. On the off chance that the wellbeing identification framework is intended to associate with compact gadgets, for example, cell phone and tablets, and so forth Residents approach certain versatile specialized instruments which are getting open. The medical services area has made patient consideration more productive. Clinical information is gathered and reported continuously detail and will be checked by specialists/medical services experts by means of PCs. Thusly, in this work a model dependent on a fake neural organization (ANN) is proposed to tackle the arrangement and forecast issue and to build up a pig influenza expectation framework that decreases the expense of real tests and furthermore the time.

## 2 Review of Literature

Saroj Kr. Biswas, Barnana Baruah, Biswajit Purkayastha, Manomita Chakraborty [01], machine learning innovation adds new potential to clinical analytic frameworks. This article presents a demonstrative model of pig influenza dependent on a fake neural organization (ANN). The proposed model chooses highlights significant for the determination of pig influenza utilizing a component choice calculation utilizing the k-Nearest Neighbor (k-NN) classifier, which diminishes the size of the information that will be utilized to prepare the ANN model. To make preparing more proficient, what is more, precise. ANN decides a slice off worth to recognize positive and negative cases, and the model positions experiments as sure or negative dependent on the cut-off esteem. The outcomes acquired with the proposed model exhibit the capacity of the model to give an undeniable degree of exactness for the analysis of pig influenza. The limit of assessment (order) of the proposed model dependent on ANN is contrasted and that of the methodologies of case-based thinking (CBR) and it is seen that the proposed model is better than the others.

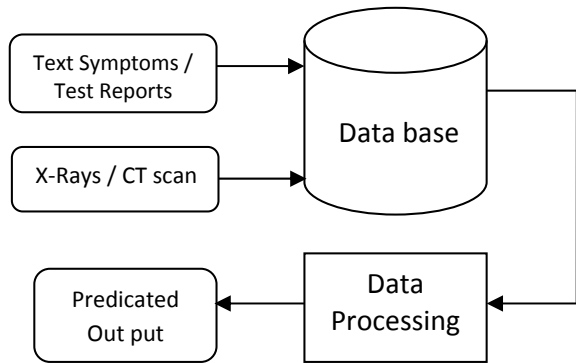
Amit Tate, Ujwala Gavhane, Jayanand Pawar, Bajrang Rajpurohit, Gopal B. Deshmukh [02], In this article, an infection expectation framework was proposed utilizing the Random Forest Algorithm (RFA). The preparation dataset is utilized to foresee a specific infection. The fundamental objective of this article is to foresee the infection which input indications are taken by the patient or by the client. Suggest a specialist who works in a specific infection if the outcome is positive. Our calculations can be reached out to work with portable/online answers for assist patients with welling for clinical finding. As an initial step, we additionally create web interfaces to assist patients with computing the danger level for every clinical case.

In this task [03], the specialists utilized different profound learning models to anticipate the probability that an individual got the H1N1 immunization or occasional influenza antibody. The specialists looked at the aftereffects of different AI models to improve the exactness of the expectation. Alongside calculated relapse, arbitrary timberland, choice trees, and so forth, we tried inclination increase calculations. The scientists ought to have anticipated if an individual will be inoculated next season dependent on their past conduct. The analysts utilized 9 calculations in the informational collection gave and had the option to accomplish a decent exactness rate with a considerable lot of them.

Ms. Ankita R. Borkar, Dr. Prashant R. Deshmukh [04], this article portrays a model that utilizes information mining procedures, called the Naïve Bayes Classifier. Information mining is an arising research pattern that helps discover exact arrangements in numerous fields. This paper features the different information extraction methods used to foresee pig influenza infections. Pig influenza is a respiratory illness that should be looked for by the patient to recognize a sickness. Progressed information mining methods help us cure the present circumstance.

Saurabh Parmar, Nihar ShahMegha Kasarwala, Madhavika Virpura, Dharmeshkumar D. Prajapati [05], From the previously mentioned enlightening review it tends to be realized that pig influenza is a hazardous illness that is spreading everywhere on the world and this is a fortuitous event thinking about that an ever increasing number of individuals in India are influenced and the cases can increment. Accordingly, it is critical to know about this sickness, as it tends to be lethal. Also, consequently, the power of this problem can be diminished by diagnosing and taking fitting medicines.

Nidhi Chauhan, Jagriti Narang, Shikha Pundir, Sandeep Singh and C. S. Pundir [07], It was distinguished that HA and NA proteins whose mix of H1N1 was of most prominent concern. On the off chance that the conclusion of a viral sickness is essentially looked for and made rapidly, pointless anti-microbial treatment can be kept away from or, whenever started experimentally, the term of anti-microbial use can be hindered and abbreviated. Quick tests for flu infections additionally permit a clinic to recognize and confine tainted patients from different patients to diminish the spread of disease. Up until now, this test is finished utilizing viral societies, quick antigen tests, and RT-PCR, yet they all have impediments. A flu biosensor might be a superior choice because of its ease, speed and high affectability factors.

**Fig. 1** proposed system

### 3 Detection System

Figure 1 underneath shows the engineering of the pig influenza discovery framework. Three phases are appeared in this chart: the principal stage is a client input stage, the subsequent stage is the back-end handling or otherwise called an information analyzer, and the third stage is the figure yield. In the principal stage, which is the client input stage, the framework will take any contribution from the client, here for our reference we are taking two sorts of info, the originally input incorporates text manifestations which are for the most part as text and another. The information is a picture, for example, X-beams and CT examines. Another stage is the information analyzer, in the wake of taking the info, it is conveyed to the dataset handling framework and afterward for grouping, where an arrangement calculation is by and large utilized for text order and here and there the SVM strategy for picture characterization. The third stage, which is the last stage, is the yield phase of the expectation which shows the forecast of the potential results.

## 4 Classification Methods

### 4.1 C5.0 Classification Method

To start with, how about we take the models gave as the premise of the choice tree. Second, we ascertain the data acquire list for each property of the preparation tests and select the attribute with the most elevated data acquire record as the division trademark. By then, we made a hub for the nature of the split and utilized the component to make a pointer for the hub. Third, we make a branch for every division characteristic gauge and as needs be partition the dataset into certain subsets. Plainly, the quantity of subsets increments to the quality number of the test include. By then, we rehash the above strides for every subset until everything subsets can fulfill one of three related conditions:



- (1) Each example in the subset has a place with a class, so simultaneously we make sheets for the subsets.
- (2) All credits that have a place with the subset have been traded and there are no unfilled ascribes left that can be utilized to isolate the dataset. At that point, in this condition, we order the subsets as the subset of classes to which the vast majority of the examples have a place and followed by an authorization made for the subset.
- (3) All test credits to one side of the examples in the subset have a similar worth, yet the examples have a place with other various classes. In this condition, we make a permit whose class is most inspected in the subset it has a place with.

Algorithm:

- (1) Create a root node for generating tree
- (2) Check the base case
- (3) Apply Feature Selection technique for analysis, Then Best Tree = Construct a decision tree using training data,
  - (A) Apply Cross-validation technique.
  - (B) Using Training set, Compute the decision tree, decide the performance accuracy with the use of Test set
  - (C) Reckon the N-fold cross-validation technique to estimate the Performance.
- (4) Apply Error reducing technique to get attribute with highest gain, Then Classification [1].

## 4.2 ANN Classification

ANN plays out an ensuing multiplication to set up the system by beginning the neurons in a secret layer. This movement starts through the BPN preparing measure utilizing preparing data gathering. The regressive spread estimation incorporates one pass forward and one pass back. The inspiration driving the forward hop is to acquire the trigger worth and the retrogressive section is to alter the heaps and inclinations as shown by the qualification among ideal and real authoritative exhibitions. These two stages will go through the experience iteratively until the organization joins. The preparation of the criticism network utilizing the back engendering calculation can be summed up as follows [2].

1. for each training pattern (presented in random order):
  - Apply inputs.
  - Summarize the output for every single neuron from the input layer, till the hidden layer and to the output layer.
  - Find the errors and calculate it at the outputs.

- Make use of the output errors in order to calculate error signals from the pre output layers.
  - For calculating the weight adjustments, use error signals.
  - Apply weight adjustments.
2. Periodically evaluate the network performance.
  3. After the training we test whether the review is positive or negative [3].

### 4.3 Support Vector Machine Classification

Backing vector machine is calculation have non-likelihood nature, which is utilized to isolate information nonlinearly and straightly. Here, dataset  $D = \{X_i, Y_i\}$  where  $X_i$  is set of tuples and  $Y_i$  is related class name of tuples. Class names are  $-1$  and  $+1$  for no and yes classification individually. The objective of SVM is to isolate negative and positive preparing model by discovering  $n-1$  hyper plane [4].

Quadratic Programming (QP) issue is ought to have been tended to in direct data. This issue is changed using the Lagrange Multipliers theory and Optimal Lagrange coefficients sets are procured. A disengaging hyper plane is made as:

$$W * X + b = 0$$

where  $W = \{w_1, w_2, w_3, \dots, w_n\}$ , where  $W_n$  is weight vector of  $n$  ascribes and  $b$  is predisposition. Distance from isolating hyper plane to any point on  $H1$  is  $1/|W|$  and Distance from isolating hyper plane to any point on  $H2$  is  $1/|W|$  so greatest edge is  $2/|W|$ . The MMH is revamped as the choice limit as indicated by Lagrangian plan.

$$D(X^T) = \sum_{i=1}^n y_i a_i x_i x^T + b_0$$

where  $X^T$  is test tuple,  $a_i$  and  $b_0$  are numeric boundaries,  $y_i$  is class mark of help vector  $X_i$ . So if sign is positive of MMH condition then XT comes in sure classification. Assuming sign is negative of MMH condition, XT comes in negative classification. SVM classifier recipe is characterized as following.

$$f(x) = \sum_{i=1}^n a_i k(x, x_i) + b$$

#### 4.4 Random Forest Algorithms

The Random Forest calculation is a set-regulated learning strategy utilized as an information indicator for grouping and relapse. In the arrangement cycle calculation, assemble a bunch of choice trees at preparing time and construct the class which is the yield method of the classes utilizing each tree. The Random Forest calculation is a gathering of tree indicators wherein each tree dependent on the estimations of an irregular vector freely experienced equivalent appropriation for all trees in the timberland. The essential rule is that a gathering of “frail understudies” can meet up to shape a “solid understudy.” Arbitrary woods are an ideal instrument for making forecasts considering they don’t fit too intently. Introducing the specific kind of arbitrariness makes them precise classifiers and relapses. Extraordinary choice trees frequently have enormous difference or high slant. Irregular Forest endeavors to direct high-change, high-predisposition issues by averaging to track down a characteristic harmony between the two limits. Taking into account that arbitrary timberlands have not many boundaries to change and can be utilized essentially with default boundary settings, they are a basic device to use without having a model or to deliver a sensible model rapidly and proficiently. Arbitrary Forests produce different orders for explicit trees. Each tree is developed as follows:

1. In the event that the quantity of conditions in the preparation dataset is  $D$ , example cases  $D$  in an arbitrary, however, with substitution state from the first dataset. This example test set will be the preparation set for developing the tree.
2. If there are input factors from the preparation dataset, a number are demonstrated to such an extent that in every hub of the tree,  $m$  factors are accessible haphazardly for it and the best division of these is utilized to isolate the hub. The estimation of is utilized as a steady all through the development of the timberland.
3. Each tree is developed to the biggest size as could be expected. There is no pruning a general adult tree. The irregular woods calculation is a set arrangement calculation dependent on the choice tree model. Produce  $k$  various subsets of preparing information from a unique informational collection utilizing a bootstrap examining approach, at that point  $k$  choice trees are worked via preparing these subsets. At last, an arbitrary timberland is worked from these choice trees. All choice trees foresee each example in the test dataset, and the last positioning outcome is returned dependent on the evaluations of these trees [5].

## 5 Conclusion

Effectively transmittable sicknesses are a significant worry for any nation’s administration and wellbeing divisions. With the coming of data innovation, it is feasible to control numerous diseases in the best and effective manner. This paper clarifies the design or strategy for foreseeing and forestalling airborne H1N1 flu utilizing IOT

procedure and AI. The C5.0 calculation with ANN is utilized to characterize clients, and SVM is utilized for proficient examination of data for AI. The central issue of the paper is the utilization of IOT strategies to look for each client contaminated with H1N1 influenza at a quicker rate.

## References

1. Pandya R et al (2015) C 5.0 algorithm to improved decision tree with feature selection and reduced error pruning. *Int J Comput Appl* 117:18–21
2. Biswas SK et al (2015) An ANN based classification algorithm for swine flu diagnosis. *Int J Know Based Comput Syst* 3:31–42
3. Borkar AR et al (2015) Naïve bayes classifier for prediction of swine flu disease. *Int J Adv Res Comput Sci Softw Eng* 5(4):120–123
4. Tate A et al (2017) Prediction of dengue, diabetes and swine flu using random forest classification algorithm. *Int Res J Eng Technol* 04:685–690
5. Parmar S, Shah N, Kasarawala M, Virpura M, Prajapati DD (2011) A review on swine flu. *JPSBR* 1(1)
6. The Merck Veterinary Manual. 2008. ISBN 978-1-4421-6742-1. Archived from the original on March 4, 2016. Retrieved 30 Apr 2009
7. Kaur EA, Gurmanikkaur (2019) Development of a computer-aided system for swine flu prediction using computational intelligence. *IRJET* 06:4818–4824
8. Sandhu R, Gill HK, Sood S (2016) Smart monitoring and controlling of pandemic influenza A (H1N1) using cloud computing. *Elsevier* 12:11–22
9. Singh V et al (2013) Pulmonary sequelae in a patient recovered from swine flu. *Lung India* 29:277–280
10. Lokwani P et al (2011) Swine flu: an overview. *J Appl Pharm Sci* 01:29–34
11. Kadam M et al (2015) A review article on swine flu. *Curr Res Pharm Sci* 05:12–16
12. Mujariya R et al (2011) A review on study of swine flu. *Indo-Global Res Lib* 1:47–51
13. Borkar AR et al (2015) Naïve bayes classifier for prediction of swine flu disease. *Int J Adv Res Comput Sci Softw Eng* 5:120–123
14. Dandagi GL et al (2011) An insight into the swine-influenza A (H1N1) virus infection in humans. *Lung India* 28:35–38
15. Thakkar BA, et al (2010) Health care decision support system for swine flu prediction using Naïve bayes classifier. In: *International conference on advances in recent technologies in communication and computing*. Kottayam, pp 101–105
16. Singh N (2013) A review article on swine flu. *Pharmatutor*
17. Chauhan N, Narang J, Pundir S, Sing S, Pundir CS (2013) Laboratory diagnosis of swine flu: a review. *Artif Cells Nano Med Biotechnol* 41:189–195. ISSN: 2169-1401
18. Patil N, Lathi R, Chitre V (2012) Comparison of C5.0 & CART Classification algorithms using pruning technique. *Int J Eng Res Technol (IJERT)* 1(4). ISSN: 2278-0181
19. Arwinderkaur, Gurmanikkaur (2019) Development of a computer-aided system for swine-flu prediction using computational intelligence. *Int Res J Eng Technol* 06(05). ISSN: 2395-0056
20. Su-lin P, Ji-zhang G (2009) C5.0 classification algorithm and application on individual credit evaluation of banks. *Sci Direct* 29(12)
21. Pandya R, Pandya J (2015) C5.0 algorithm to improved decision tree with feature selection and reduced error pruning. *Int J Comput Appl* (0975–8887) 117(16)
22. Sandhu R, Gill HK, Sood SK (2015) Smart monitoring and controlling of pandemic influenza A (H1N1) using social network analysis and cloud computing. *J Comput Sci*
23. Gandhi S (2018) Support vector machine—introduction to machine learning algorithms. SVM model from scratch

24. Zurada JM (1992) Introduction to artificial neural network. PWS Publishing Company
25. Zhang GP (2000) Neural networks for classification: a survey. IEEE Trans Syst Man Cyber Part C: Appl Rev 30(4):451–462

# Text Mining and Machine Learning Algorithms to Identifying Diseases and Providing Repair Action Using ICD-10 Codes



Ashish P. Ramdasi  and S. Sathyalakshmi

**Abstract** As we know that there are many doctors in all around the world, each doctor treats their patients based on their knowledge and skills. So there may be a possibility of misdiagnosis or sometimes rare diseases may be unnoticed by the doctors. It is really very hectic process for most of the doctors as they have to provide ICD-10 CM code after diagnosis. Hence, we have proposed a predictive system to identify diseases based on symptoms and give best treatment to the patient using text mining and machine learning algorithms. Our system identifies the generalized diseases and provides treatment. Our system is helpful for those patients who cannot afford to go to the physician for treatment. Also, our system is useful for doctors as it makes their work easy as we will provide ICD-10 CM code after classifying the disease. Here, in our system, the patient provides unstructured data as an input to the system and the system will determine diseases from the data provided by the system and suggest the best suitable drugs to treat their diseases. Our models consist of classifiers such as Random Forest, Naive Bayes, Support Vector Machine and Logistic Regression machine learning algorithms. Based on the performance of the classifiers we are doing the comparative analysis to know which classifier is more accurate based on their accuracy score.

**Keywords** Text mining · ICD-10 CM code · Unstructured data · Machine learning · Random Forest · Naive Bayes · Generalized diseases

## 1 Introduction

There are many innovative and significant machine learning application in medical field. Everyday huge amount of medical data is created through doctor's notes, electronic medical data, prescriptions and other sources. This data can be put to some

---

A. P. Ramdasi (✉) · S. Sathyalakshmi  
Hindustan Institute of Technology and Science, 1, Rajiv Gandhi Salai (OMR), Padur,  
Kelambakkam, Chennai 603103, India

S. Sathyalakshmi  
e-mail: [slakshmi@hindustanuniv.ac.in](mailto:slakshmi@hindustanuniv.ac.in)

efficient use such as to predict the diseases. There are many systems created to detect diseases such as in Sudeshna et al. [1] proposed an approach to identify symptoms and treatment for Heart Diseases [2], which is useful for those patients having heart problems; this system is not useful for detecting generalized Diseases. Our system is used to detect generalized diseases, it is useful to all users who will and will not be able to afford expensive visits to hospitals and is not able to take appointment of doctors as they cannot afford doctor’s appointment fees. In addition, it is also useful for doctors to detect the disease easily, as different doctors having different knowledge will provide different treatment to the patients based on their experience and knowledge but sometimes doctor’s will ignore simple symptoms which will later cause serious disease to the patient. Our system will predict most accurate disease based on symptoms provided by patient. Our system will take maximum number of symptoms as input to increase accuracy and so that serious disease will be detected at early stage. Our paper also contributes to help doctor to identify the ICD-10 CM code [3] of the particular disease easily.

Here in our system we used Machine learning and text mining techniques to predict diseases.

1.1 Architecture

The main aim of the architecture of our system is to predict the most accurate disease and provide suitable treatment to the diseases (Fig. 1).

- (1) *User*: The user will enter the symptoms and these symptoms will be taken as input for identifying diseases and providing treatment from symptoms using text mining in [6–8] and machine learning algorithms [9–12] such as Random Forest and Naive Bayes.

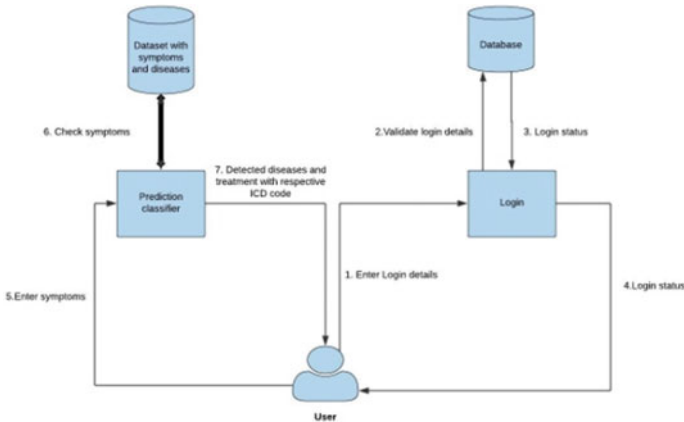


Fig. 1 System architecture [4, 5]

- (2) *Prediction Classifier*: Here, the models used are Random Forest and Naive Bayes. Using these models, it will help to analyze the diseases from symptoms. These models will use patient data to accurately identify the diseases [13, 14] and our system will be provided with unstructured data entered by the patient and our model will extract symptoms from that unstructured text and will further predict diseases
- (3) *Database*: There are two databases in our system one dataset include symptoms, diseases, treatment and ICD code which is used by our predictive models to detect diseases and other database will include login details of patient which patient entered after registering in our system.
- (4) *Treatment Identification and providing ICD code*: After detecting diseases, then according to the diseases it will also recommend treatment.

## 2 Existing System

In the traditional architecture, there existed only the server and the client. In most cases, the server was only a database server that can only offer data. Every client has to be trained as to use the application and even security in the communication is also the factor to be considered in Kohli and Arora [15]. The current system consists of a Medical Data Warehouse. This consists of all the details of the patients approached for the diagnosis. The medical history of the patient is recorded in the system. This data allocation in the system is done by the Doctor. The Doctor refers the diagnostic document uploaded by the patient. The diagnostic document is forwarded further to the Operator. The Operator associates a unique ICD-10-CM code to every disease and the respected data is forwarded to the Accounting Officer for the inspection. The diagnostic document is rechecked by the Accountant and the result is recorded, respectively. The same is transferred to the Medical Data Warehouse for the data retrieval, respectively. The issue occurred in any of the body part is determined in the existing system. The existing system goals disease detection [16–18] in the whole body and the recovery (Fig. 2).

## 3 Proposed System

The patient history is in the form of text data these sentences are filtered and the keywords (meaning/output correlated words) are extracted using BoW technique. TF-IDF approach is also used while preprocessing. Finally model build using classifier to identify the disease encountered in any of the body part and the repair action is provided to patient.



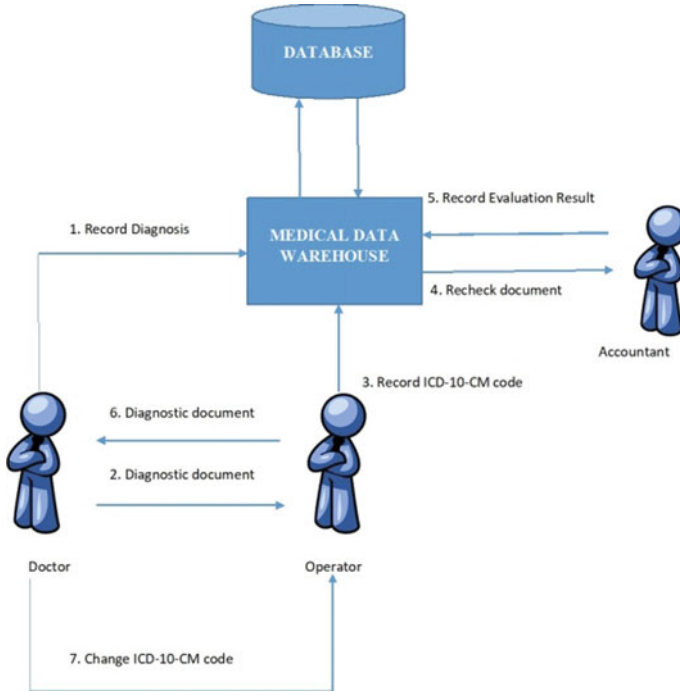


Fig. 2 Existing system [4, 5]

### 3.1 Collecting the Data

All the data collected includes the details of the history of the patient. The source of the collection of data is Nobel Hospital and the Ruby Hall Clinic. The ICD-10-CM code is associated with the data which is collected from both of the sources.

### 3.2 Building the Model

The procedure of the building of the model is categorized into two parts: Data preprocessing and the Use of the classifiers in the classification.

- (1) *Data preprocessing*: In this part, the unstructured text is provided as the input to the system. The sentence is filtered and the keywords which describe the uneasiness in the body are taken out of the sentence and further forwarded to the classifier for the accurate identification of the disease [19].

This technique named as the Bag of Words is used to tokenize the sentence and the required words are extracted from the sentence. This process removes

all the noisy data from the sentence and the sentence gets filtered for the further process.

- (2) *Use of the classifier in the classification process:* In this system two classifiers are mainly used for the identification of the disease. They are termed as Nave Bayes classifier and the Random Forest classifier [20]. The words depicting the uneasiness in any of the body part are extracted by the classifiers and the result is generated, respectively.

## 4 Classifying Methods

The algorithms which we are going to implement are Random Forest and Naive Bayes for classifying diseases from symptoms in [21, 22]. Random Forest and Naive Bayes are the type of supervised learning algorithms.

Gaussian Naïve Bayes [23, 24] supports continuous value data and follows gaussian normal distribution. To follow this approach, the data must be described by Gaussian distribution with no co-variance (independent variable). By calculating mean and standard deviation of each label is taken into consideration while fitting the model. At every data point, Z-score distance between that point and each class-mean is finding-out. It will improve the accuracy.

Classification Steps:

- 1: Import datasets where diseases are stored in separate variable and given unique code, to make our classification easier.
- 2: Split training and testing data.
- 3: Train the model with the help of any of above classifier model.
- 4: We need to fit the dataset in this model.
- 5: We perform prediction [25, 26] and based on predicted value we will check the disease associated with the predicted unique code.
- 6: Based on predicted disease, the treatment is extracted for that disease.
- 7: Evaluate accuracy score of predictions.

## 5 Dataset Description

Our sample data consists of 4921 medical records. Our dataset consists of two Comma Separated Values (CSV) files and two excel sheets. Training.csv and Testing.csv consists of attributes such as symptoms and prognosis, treatment.xlsx consists of attributes such as disease and treatment used to provide treatment to patients and icd.xlsx consists of attributes such as diseases and ICD-10 CM CODE used to provide ICD code to doctors. Our data consist of 40 diseases and 132 symptoms. We take Training.csv and Testing.csv as input to the system and diseases are classified according to data in these files.

## 6 Experimental Results

We measured the performance of our models using the training and testing datasets. Our models Random Forest and Naive Bayes are compared based on their accuracy score, precision score, F1 score and recall. Confusion matrix is used for performance measurement.

- A. *Accuracy*: Accuracy is percentage that our classifier predicts correctly.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- B. *Precision*: Precision designates the percentage of your results which are pertinent.

$$\text{Precision} = \frac{TP}{TP + FP}$$

- C. *Recall*: Recall refers to the percentage of total pertinent results rightly relegated by your algorithm.

$$\text{Recall} = \frac{TP}{TP + FN}$$

- D. *F1 score*: F1-score, is harmonic mean of precision and recall.

$$\text{F - measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

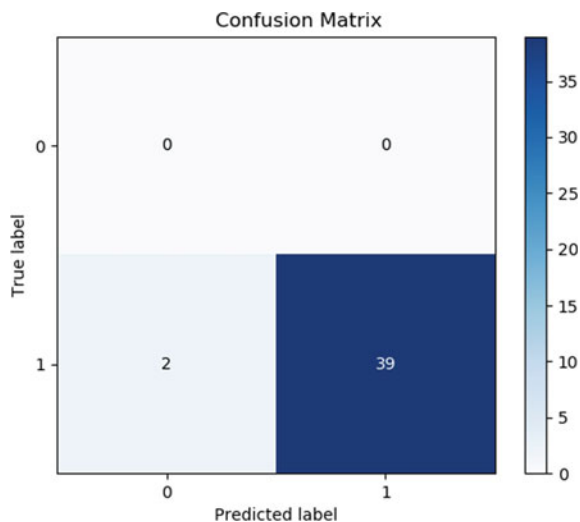
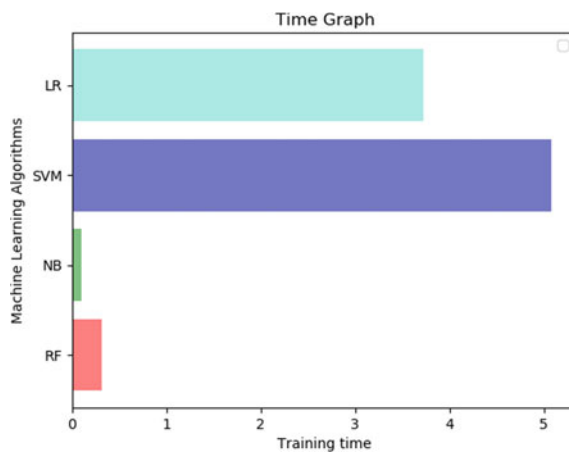
- A. *Confusion Matrix*:

A confusion matrix [27] is used to describe the performance of a classifier on testing data for which the true values are known (Figs. 3 and 4; Table 1).

By referring to this graph, it is clear that Support Vector Machine Algorithm takes much longer time for prediction which is 5.310 s. Naive Bayes algorithm takes 0.115 s which is least time as compared to other algorithms.

## 7 Conclusion

Our system helps to predict the accurate disease without any misdiagnosis. Random Forest, Naive Bayes, Logistic Regression and Support Vector Machine algorithm is proposed in this paper to improve the efficiency and accuracy for prediction. We recognize the symptoms and based on the symptoms the disease is classified and based on the patient dataset we identify the best drug that would set for their health conditions. Random Forest yields accuracy of 95.

**Fig. 3** Confusion matrix [4, 5]**Fig. 4** Time graph [4, 5]**Table 1** Results [4, 5]

	Precision	Recall	F1 score	Accuracy
Random Forest	92.6829	95.1219	93.4959	95.121
Naive Bayes	91.6829	94.1219	92.4959	93.12
LR	92.00	95.00	97.00	95.121
SVM	92.00	95.00	97.00	95.00

Furthermore, new classifying methods can be developed by combining these algorithms to get a broader perception of the significant features to increase the performance of disease prediction. Further, the project can be extended by classifying the disease based on seasons, and predict the disease more accurately by considering the current season.

**Acknowledgements** We would express our gratitude to the HOD(CSE),HITS, Chennai for giving us the necessary permission and sanction to use all resources for this research. Thanks to all the staff and faculty members of HITS, Chennai for giving us necessary guidance in completing this research work. Sinhgad Academy of Engineering also gave the necessary permission in completing the research work.

## References

1. Sudeshna P, Bhanumathi S, Anish Hamlin MR (2017) Identifying symptoms and treatment for heart disease from biomedical literature using text data mining
2. Gandhi M, Singh SNF (2015) Predictions in heart disease using techniques of data mining
3. Xie S, Fan H (2019) Research on CNN to feature extraction in diseases prediction
4. Burkul P, Umapathy K, Asaithambi A, Huang H (2020, September) Data mining pipeline for performing decision tree analysis on mortality dataset with ICD-10 codes. In: Proceedings of the southern association for information systems conference, p 1
5. Fabacher T, Godet J, Klein D, Velten M, Jegu J (2020) Machine learning application for incident prostate adenocarcinomas automatic registration in a French regional cancer registry. *Int J Med Inform* 139:104139
6. Ketpupong P, Piromsopa K (2018) Applying text mining for classifying disease from symptoms. In: International symposium on communications and information technologies
7. Aich S, Sain M, Park J, Choi KW, Kim HC (2017) A text mining approach to identify the relationship between gait-Parkinsons disease (PD) from PD based research articles
8. Erraguntla M, Gopal B, Ramachandran S, Mayer R (2012) Inference of missing ICD 9 codes using text mining and nearest neighbor techniques
9. Dahiwade D, Patle G, Meshram E (2019) Designing disease prediction model using machine learning approach. In: Third international conference on computing methodologies and communication
10. Mohan S, Thirumalai CS, Srivastava G (2019) Effective heart disease prediction using hybrid machine learning techniques
11. Ramesh S, Vydeki D (2018) Rice blast disease detection and classification using machine learning algorithm
12. Escudero J, Ifeakor E, Zajicek JP, Green C (2018) Shearer, machine learning-based method for personalized and cost-effective detection of Alzheimer's disease
13. Chen M, Hao Y, Hwang K, Wang L, Wang L (2017) Disease prediction by machine learning over big data from healthcare communities
14. Nishara Banu MA, Gomathy B (2014) Disease forecasting system using data mining methods
15. Kohli PS, Arora S (2018) Application of machine learning in disease prediction
16. Duangchaemkarn K, Chaovatur V, Wiwatanadate P, Boonchieng E (2017) Symptom-based data preprocessing for the detection of disease outbreak
17. Huang F, Wang S, Chan CC (2012) Predicting disease by using data mining based on healthcare information system. In: IEEE international conference on granular computing
18. Ed-Daoudy A, Maalmi K (2019) Real-time machine learning for early detection of heart disease using big data approach

19. Kurdia H, AlMansour N (2017) Identifying accurate classifier models for a text-based MERS-CoV dataset
20. Patil M, Lobo VB, Puranik P, Pawaskar A, Pai A, Mishra R (2018) A proposed model for lifestyle disease prediction using support vector machine
21. Prabakaran N, Kannadasan R (2018) Prediction of cardiac disease based on patients symptoms. In: 2nd international conference on inventive communication and computational technologies
22. Shankar M, Pahadia M, Srivastava D, Ashwin TSG, Reddy RM (2015) A novel method for disease recognition and cure time prediction based on symptoms second international conference on advances in computing and communication engineering
23. Shinde R, Arjun S, Patil P, Waghmare J (2015) An intelligent heart disease prediction system using k-means clustering and nave bayes algorithm
24. Kaur S, Kalra S (2016) Disease prediction using hybrid K-means and support vector machine
25. Kohli PS, Kohli PS (2018) Application of machine learning in disease prediction
26. Repaka AN, Ravikanti SD, Franklin RG (2019) design and implementing heart disease prediction using Naives Bayesian
27. Ramdasi AP, Mehata KM (2019) Improved text mining algorithm for fault detection using combined D-matrix

# Communication Skills: An Integral Instrument for Data Analysis in the Field of Research



Shweta Arora , Kavita Ajay Joshi , Preet Nain Kaur Narula ,  
Deepa Nainwal , Narendra Bisht , Girish Nainwal ,  
and Ayush Agarwal

**Abstract** There is an increasing demand of high-quality research in every field. Researchers are exploring varied domains and serving the society. It has often been observed that the researchers are excelling in their core area but due to lack of effective communication skills they are unable to present the data of their research findings in an effective and desired manner which eventually impedes the optimum exposure and utilization of their credible research work. This emphasizes on the fact that communication skills not only add value to the research but also acts as a spinal cord of the research. The present paper seeks to identify the gap between communication skills and data analysis in the field of research. A survey approach was used to identify and confirm the need of communication skills for data analysis in research. In the present research work, after the diagnosis of the problem, an effective counteractive phenomena which included employment of self-analytical activity -SWOC analysis, public speaking-Delivering Presentation and Impromptu Speaking, Reading Comprehension-Reading Articles and Research Papers and Technical/Professional Writing- Article, Synopsis and Research Paper Writing.

**Keywords** Data analysis · Communication skills · Research · Optimum utilization · Barriers · Interpret · SWOC · Vocabulary building

## 1 Introduction

Data analysis and communication skills go hand in hand. It is important for the data analysts/scientist/researchers to have an expertise in the technical domain as well as in communication skills. The skills required by researchers/analysts should be diverse and wide-ranging from analysis, technical knack, domain specific information to effective communication skills [1]. Various other studies have also inferred that data analysts must utilize a combination of skillset that includes technical as well as

---

S. Arora (✉) · K. A. Joshi · P. N. K. Narula · D. Nainwal · N. Bisht · G. Nainwal · A. Agarwal  
Graphic Era Hill University, Bhimtal, Uttarakhand, India

soft/non-technical skills [2, 3]. Data Science generally deals with the visualization of models and figures derived from some data collected by the researchers/data analyst. The models created are then presented to the audience. To support the research approach in interpreting data, expressing findings, and providing recommendations effectively, it is important to integrate communication skills with data analysis in research. According to Bagozzi, the process of theory formation, after interpreting the data is also important [4, 5]. The models formed will have little value if they are not expressed sensibly. Jyotsna Vadakkanmar has rightly quoted from Neha Shitut's article that "analytics is not only about numbers, but it is also about making sense of the numbers" [6]. So, it is important for a good researcher to decipher the output of his research and convey it effectively to the audience. Communication skills including soft skills play an indispensable role in expressing the data effectively and rationally [7]. Effective communication skills facilitate in communicating the findings in an influential manner. But it is unfortunate that the research scholars in large numbers failed to express their research observations and findings due to the lack of good communication skills both written and oral. This resulted in the rejection of manuscripts or getting revision from the reviewers of reputed journals. It was observed that the research scholars basically focused on the technical domain-from collection of data to its interpretation and visualization. But in the entire process the unseen thread of effective communication skills was overlooked. This at times resulted in the misinterpretation of data as well. The existing problem was examined in this study. The gap between data science and communication skills needed to be bridged. To bring an amicable change in the analytical setting, there was a need to enhance the non-technical skills [8] of technical researchers. So, there was a need to stipulate Remedial English classes to the research scholars before they penned down the observations and findings of their study. The present paper aims to discuss the application and results of the techniques employed in the Remedial English Communication classes to get an effective research outcome. According to many scholars, researchers from various fields have turned to "literary and textual analysis" for a better explanation of the data [9, 10]. This is a positive insinuation for importance of communication, i.e., written and oral in data analysis. Also, it demands proficiency in choosing apt vocabulary and sentence construction so that the new research work is utilized by the community at large for deriving maximum benefit [11, 12].

To improve the quality of research, it was a prerequisite to identify the gaps and incorporate training modules of fundamentals of grammar, vocabulary building, and presentation skills, etc., to data analysis [13, 14].

## 2 Research Tools and Methodology

An empirical study was employed with the following details:

Sample size: 30 Research Scholars



### Educational Background: Masters Students of Technical Background

The Following tools were adopted:

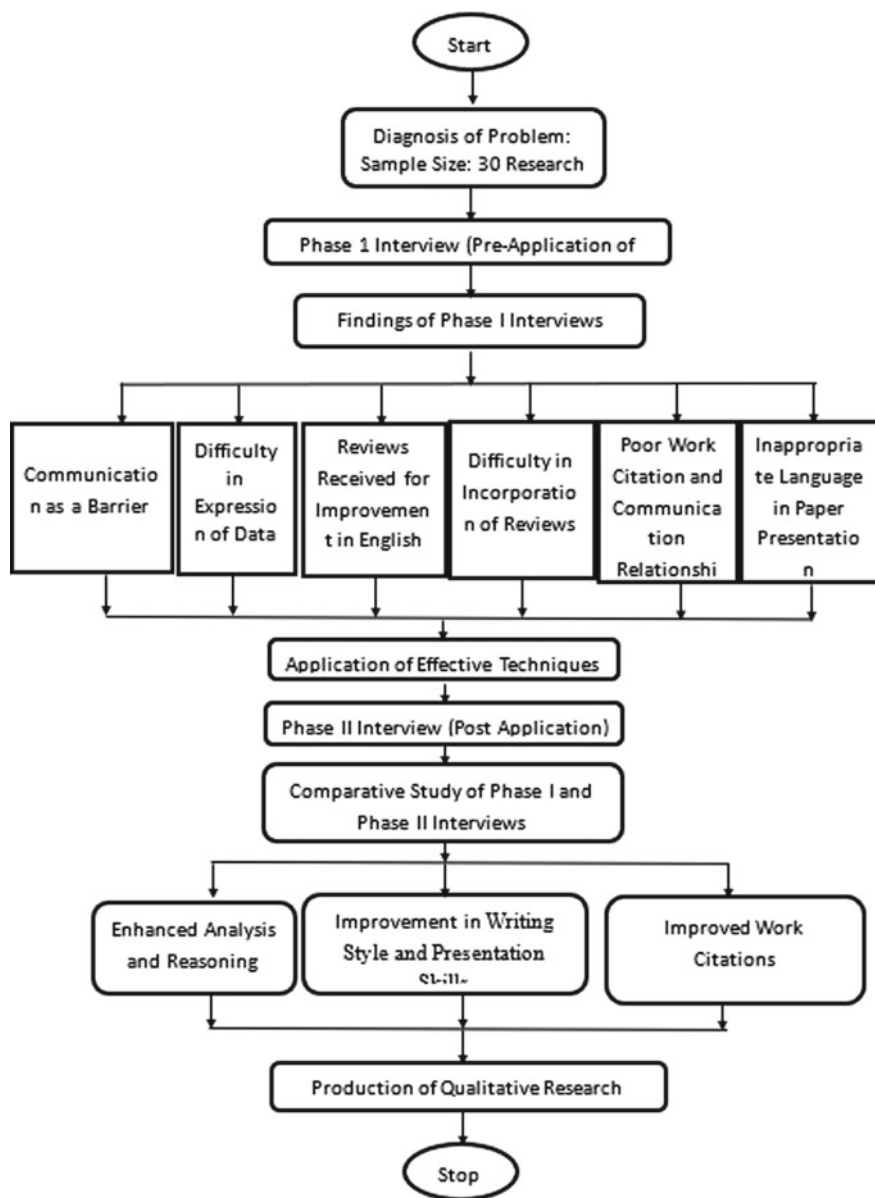
1. Online Interviews to gauge the intensity of the problem.
2. Participatory Drills-SWOC, Visual Stimuli, Delivering Presentations, Articulation.
3. Verbal Ability Tests-Reading Comprehension, Fundamentals of Grammar and Advanced Vocabulary (Fig. 1).

## 3 Action Plan

During covid-19 pandemic all interactions were carried out in a virtual platform. Therefore, online interviews of the research sample (30 research scholars) were held on MS Teams platform. It was carried out in two phases using a questionnaire that included questions related to the topics like importance research in an organization, importance of effective data analysis in research, importance of oral and written communication in research, problems encountered by researchers in portraying their findings, the role of soft skills in data reorientation, etc. Each interview lasted for 15 to 20 min. The first phase of interviews gave a better understanding of the problems faced by the research scholars. Whereas the second phase gave the feedback, or the results of the communication strategies applied to their technical research.

1. Is communication skill a barrier to effective research?
2. Do you find it difficult to express your research data in English language?
3. Have you received reviews related to improvement of English Language for the revision of your research papers?
4. Do you find communication skills as a major challenge in improving the paper according to the comments of the reviewers?
5. Is there any relationship between research work citation and communication skills?
6. Do you think you need apt language to present your paper in research seminars and conferences?

After interviewing the research scholars in Phase I, it was discovered that communication was a great barrier in the process of research paper writing. The major difficulty the researchers faced was the expression of data interpretation in their research findings. Poor data analysis resulted into more rejections of the research paper which led to a negative impact on the minds of the researchers. Also, among various reviewers' comments on a paper poor communication being predominant. Researchers also found that there was a great difficulty on working on improvements in English to get their work accepted in good journals. The researchers found that it was required to have good communication skills along with apt language to produce quality research work. Somewhere there was a direct relationship between work citation and good communication. The work written in lucid and comprehensible language was read more and there were more cited by other researchers.



**Fig. 1** Flowchart indicating diagnosis of problems and remedial action plans

After understanding the intricacies of the situation of the researchers, the findings of the above tool helped in developing an action plan for the researchers so that they could attain proficiency in the expression of their research findings in the English language. Keeping in mind the need of the hour, remedial classes were scheduled on weekends for the research scholars to enhance their communication skills. Concepts on parts of speech was delivered to the researchers so that the scholars could frame correct grammatical sentences. Certain practically relevant approaches were adopted. Practical devices like SWOC analysis, pictorial/graphical presentations, vocabulary building exercises, reading skills, presentation skills and pronunciation practices, etc., were carried out to assist data science management among researchers. Some of the procedures are deliberated below:

There was a special focus on the topics like tenses, subject-verb agreement, articles, and prepositions which would aid them in interpreting the data correctly.

Assess the fundamentals of grammar so that the research scholars could identify the area of improvement and work to acquire proficiency in the verbal domain. A series of diagnostics tests were given to strengthen the verbal area of the research scholars. A sample diagnostic test is given hereunder (Fig. 2).

There was a need to incorporate vocabulary building exercises so that the researchers could employ good synonyms and antonyms in their work. Some sample activities that were carried out to enhance vocabulary of the researchers as hereunder:

A concept session on root word technique was given to the researchers. They were provided a list of root words along with their meanings and were asked to form words of the given roots. Though such process of learning vocabulary demanded more time, yet the outcome of the activity was fruitful as it came out as an effective vocabulary retention technique. A few sample sheets of the activities are given Table 1:

SWOC analysis an introspection drill helped the research scholars to find out their attributes specifically their weaknesses and challenges. In these sessions, the research scholars were able to identify and overcome their weaknesses and challenges (Figs. 3 and 4).

Visual Stimuli activity the researchers were given to describe live objects/pictures/scenes. Such simple activities require effective cognition as the process involves conversion of thoughts into words. Some sample questions related to data analysis given to the researchers are hereunder:

According to a survey the following data was collected for “Comfort Food for Teenagers.” The pie chart shows the percentages, and the table shows the data. Now write the names of the food items on the pie chart over/near the percentages from the table given. Also fill your choices from 1–6 and comments in the “**Your Choice Comments**” column (Fig. 5).

<b>DIAGNOSTIC TEST-1</b>	
<b>Time: 40 Min</b>	<b>MM: 25</b>
<b>Q.1 Fill in the blanks with the correct prepositions:</b> <span style="float: right;">(5 marks)</span>	
i. Rahul's friend died..... chikungunya. ii. My husband loves to listen..... classical music. iii. My parents agreed.....my decision for going abroad for further studies. iv. The plane took..... for Bangalore early morning. v. Mangala caught me .....my hand to share her experience.	
<b>Q.2 Fill in the blanks with correct sub-verb agreement</b> <span style="float: right;">(5 marks)</span>	
i. Idli and Sambhar .....my favorite South Indian cuisine. (is/are) ii. The singer and the performer.....arriving for the annual fest. (is/are) iii. The cattle..... grazing in the meadows. (was/were) iv. The news of the sudden militant attack.....a great shock. (was/were) v. Either Gunjan or Hitharthi.....the winner of Sudoku. (was/were)	
<b>Q.3 Insert the correct article in the following sentences:</b> <span style="float: right;">(5 marks)</span>	
i. This is .....most beautiful painting I have ever seen. ii. ....angel appeared in my dream and showered good luck and prosperity. iii. Osheen wants to become ..... tallest girl of her class. iv. Ruchira completed her MD from..... U.S. v. ....earth is the only planet where life survives.	
<b>Q.4 Insert the correct form of adjective in the following sentences:</b>   (5 marks)	
i. Karan is ..... intelligent of the two friends. (more/most) ii. He is the ..... son of Kapoor family. (oldest/eldest) iii. Aahna is the..... girl in the campus. (tallest/taller) iv. In winters, the nights are.....than the days. (long/longer) v. Manglam is ..... of the two sisters. (cleverer/cleverest)	
<b>Q. 5 Describe your dream in 250 words. (Using appropriate tenses) (5 marks)</b>	

**Fig. 2** Diagnosis test

**Table 1** Activity on word formation through roots

Root	Meaning	Form words
Acro	Top	Acronym,.....,
Brev	Short	.....,
Cred	Believe	.....,
Derm	Skin	.....,
Ject	Throw	.....,
Pater	Father	.....,
Quasi	Resembling	.....,
Vice	Acting in place of	.....,

1. EXTROVERT:  
Root Word: .....  
Meaning: .....  
Usage: .....

2. DISPARAGE:  
Root Word: .....  
Meaning: .....  
Usage: .....

3. RIDICULOUS:  
Root Word: .....  
Meaning: .....  
Usage: .....

1. BRIEF:  
Meaning: .....  
The employees were briefed on the insurance policies.

2. ESCALATING:  
Meaning: .....  
One should not escalate petty issues.

3. EVENTUALLY:  
Meaning: .....  
The thief was eventually arrested by the police.

Fig. 3 Activity on usage of root words

STRENGTHS	WEAKNESSES
OPPORTUNITIES	CHALLENGES

Fig. 4 SWOC analysis

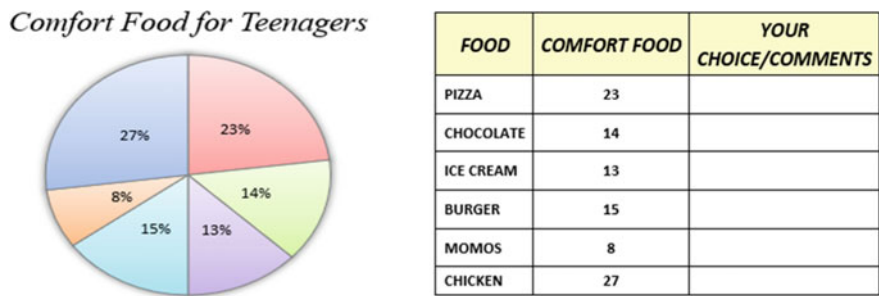


Fig. 5 Activities on visual stimuli

4 Data Analysis

A sample size of 30 research scholars was taken. The data of pre- and post-application of the techniques was observed and studied. The observations were further categorized as high, average, and low. A comparative study of pre- and post-application of techniques was done. The outcome of the analysis is deliberated in the conclusions. The analyses of pre- and post-findings of the present study reveals the effectiveness of communication skills in research (Fig. 6).

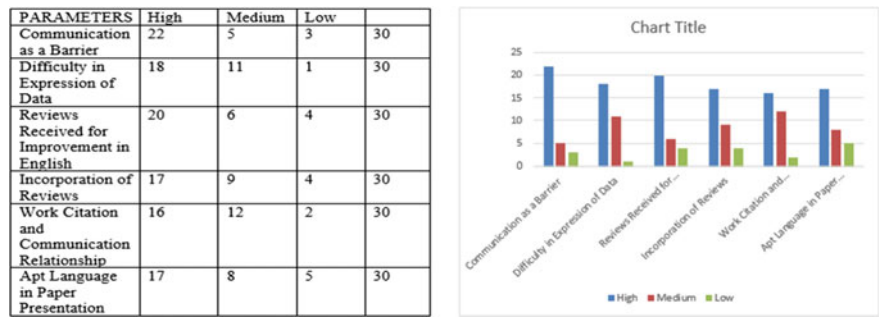


Fig. 6 Pre-application (phase I interview findings)

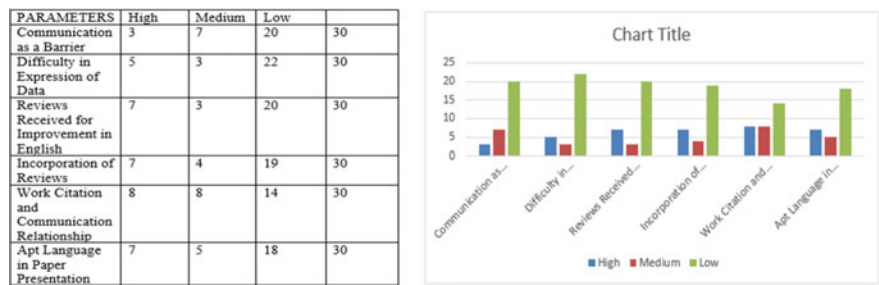


Fig. 7 Post-application (phase II interview findings)

In Phase I, 73.3% of the study sample considered communication as one of the foremost barriers in expressing themselves. This statistic fell to 10% in Phase 2. This indicates that practice session both for written and oral played a vital role in overcoming the barriers to effective research work which depends on specific analysis in their findings (Fig. 7).

Only 16% of the study sample found difficulty in expression of data. Comparing Phase 1 and Phase 2, a visible improvement of 44% was seen in this aspect. As we all know that good data analysis is a spinal cord of an effective research, the drills/ concept sessions/ test series adopted in the present study can be utilized for research scholars and Masters Students for improving their analytical skills in article/synopsis/thesis/ research paper/proposal writing.

Phase 2 result indicates that the incorporation of practice drills and grammar sessions have aided 66.6% of the study sample to have fewer reviews received for improvement in English language. However, Phase 1 analyses reveal that the same percentage of our sample size received reviews for improvements in English language. In Phase II, only 23.3% of the sample received such reviews. This remarkable improvement indicates importance of English language teaching /learning for quality research work.

The data analyses of Phase 2 findings reveal that almost 63.3% of the sample size was able to incorporate the Reviewers comment in their research papers. Such noticeable improvement will result into increase in the number of research publications in good journal and will also correspond to minimal rejections and revisions of research papers.

The result of Phase 2 related to work citation and communication relationship indicates that 73% of the study sample was with the opinion that their work published after participating in practice drill, taking test series, and attending concept sessions. Since there is a marked improvement in the quality of content writing and analysis of their findings, there was an increase in citations of their published works.

The data analysis of Phase II of this study shows that only 23.3% of the study sample lacks apt language in paper presentation, whereas 76% of the study sample have shown a drastic improvement in research paper presentation in various seminars and conferences. This was a result of consistent overhauling of personality through SWOC analysis.

## 5 Conclusions and Suggestions

The researchers confronted grammatical glitches in analysing and interpreting the data. The concept sessions on fundamentals of grammar and miscellaneous practice sheets helped research scholars improve in this area. Small composition writing helped them in being thoughtful and coherent in writing their research articles/papers.

The techniques used to enhance vocabulary had a positive impact in the performance of the research scholars. They were able to express their ideas verbally. There was a remarkable improvement seen in the writing style and presentation skills of the research scholars. With the enhanced word power, the research scholars used good vocabulary in writing their research paper/article and presented it well.

There were intrapersonal and interpersonal barriers among researchers in the expression of the idea. SWOC analysis helped the researchers to overcome these barriers effectively and they were able to present their content confidently using kinesics in paper presentation in seminars and conferences.

There were poor visual stimuli among researchers which resulted into ambiguous data interpretation and vague result output. The primary benefits of using visual stimuli technique observed among researchers were enhanced analysis and reasoning, better construal/interpretational skills, increased eloquence, improved articulatory, boosted thought process and a good command on the language for writing and presenting their ideas with respect to data analysis and research presentation.

The present study reveals that there is a subtle but direct relationship between communication skills and data analysis in the field of research. The unseen thread of communication skills helps to weave the technical findings and analysis efficiently.

## References

1. Hsinchun CRHL, HStorey VC (2012) Business intelligence and analytics: from big data to big impact. *MIS Q* 36(4):1165–1188
2. Mauro DA, Greco M, Grimaldi M, Ritala P (2017) Human resources for bigdata professions: a systematic classification of job roles and required skillsets. *Inf Process Manage* 54(5):807–817. <https://doi.org/10.1016/j.ipm.2017.05.004>
3. Debortoli S, Muller O, von Brocke J (2014) Comparing business intelligence and big data skills: a text mining study using job advertisements. *Bus Inf Syst Eng* 5:289–300. <https://doi.org/10.1007/s12599-014-0344-2>
4. Bagozzi RP (1984) A Prospectus for theory construction in marketing. *J Mark* 48(Winter):11–29
5. Spiggle S (1994) Analysis and Interpretation of qualitative data in consumer researchanalysis and interpretation of qualitative data in consumer research. *J Consum Res* 21:491–502
6. Vadakkanmar J (2015) Communication and storytelling-a crucial skills for data analyst. <https://www.jigsawacademy.com/communication-story-telling-crucial-skill-data-analysts/>
7. Mitchell RB, Woolridge RW, Johnson V (2020) The role of non technical skills in providing value in analytics-based decision culture. *J Edu Bus*. <https://doi.org/10.1080/08832323.2020.1719961>
8. Boomer L (2018) Business analytics ... new skills, new services! *Account Today* 32(7):1
9. Krupat A (1992) *Ethnocentrism*. University of California Press, Berkeley and Los Angeles
10. Marcus G, Fisher M (1986) *Anthropology as cultural critique: an experimental moment in the human sciences*. University of Chicago Press, Chicago
11. Arora S, Joshi KA, Koshy S, Tewari D (2017) Application of effective techniques in teaching learning english. *Engl Lang Teach* 10(5):193–203
12. Arora S, Joshi KA, Koshy S (2018) *Professional communication: practical workbook*. Spire Publications, India
13. Arora S, Joshi KA, Pant G, Bisht N, Sharma K (2020) Mnemonics-an effective tool for vocabulary building. *Int J Psychosoc Rehabil* 24(6):4957–4964
14. Joshi K, Arora S, Sabarwal P, Agarwal P, Gabbhir A (2020) Contemporary personality litmus test through SWOC analysis. *Test Eng Manag* 83:659–666



# Model for Intrusion Detection in Cyber-Physical System to Address Network Simulation



R. Aishwarya, M. Ajitha, and A. Sheryl Oliver

**Abstract** Network security is essential for both personal and business networks. Somewhat recently, the remarkable development of organization related administrations has brought about an enormous inundation of secret information on the web. Be that as it may, networks are exceptionally defenceless against unauthenticated interruptions. Current digital actual frameworks are getting progressively related. Such interdependencies make new weaknesses and make these frameworks more defenceless to disappointments. There are many cyber-attacks such as Denial of Service attack, Man in the Middle attack, SQL injections. The proposed system is intrusion detection system which detects for various failures and attacks in the given network data. It uses ML algorithm, for detection of the various attacks and provides the result as the same. The model has been evaluated on dataset collected through network simulations and various data collected from the online sources. The proposed system make sure the system is able to detect the intrusions effectively.

**Keywords** Network security · Cyber-physical system · Machine learning · Network attack · Intrusion detection

## 1 Introduction

A cyber-physical system may be a set of systems of various sorts whose primary objective is to observe a physical method and reply to ever-changing conditions in the period through feedback. Physical processes, networking, and computation all close to form them. Fashionable networks, like the sensible grid and communications networks, are getting more and more mutually beneficial.

The word “real” refers to the article because it is viewed by human senses, whereas “cyber” refers to a simulated illustration of the atmosphere within which the entity exists, yet as extra info regarding the article, like its build or model. Having the choice

---

R. Aishwarya (✉) · M. Ajitha · A. Sheryl Oliver  
Department of Computer Science and Engineering, St. Joseph’s College of Engineering, Chennai  
600119, India

to ascertain things from each physical and virtual viewpoint, digital actual frameworks—below human watch—will autonomously assess operational conditions and backing guaranteeing dynamic and operational variations.

Correspondence is key in digital actual frameworks, as they allow varied things to trade knowledge with each other and with folks, whenever and in any conditions. These intricate interdependencies may seriously influence network functionalities. A notable illustration of such impacts happened in Italy in 2003. These occasions are normal, as somewhere in the range of 2003 and 2012, in excess of 600 territorial blackouts happened in the USA, which influenced a great many individuals. Organization execution—as far as dormancy, transfer speed, and dependability—to a great extent impacts the cooperation between the various parts of a digital actual framework, and the capacity to execute equal assignments inside a particular time period is basic.

Another vital test for digital actual frameworks is distinguishing basic issues and blames. In a complex digital actual framework, a minor deficiency can create a course of cautions, making it hard to recognize the issue and its underlying driver. Because of the number of digital assaults against governments and business organizations all throughout the planet, network interruption identification frameworks (NIDS) have been quickly evolved in the scholarly world and industry. There are various techniques, for instance, ML to interference revelation structures which are requested into coordinated, solo, and backing learning and can develop the basic model subject to planning data. In oversight ML, both common place and irregular practices are given to the model to learn arranged named data. It is especially hard for aggressors on advanced genuine structures to get named data while they don't need irregular data in the arrangement stage and it is an uncommon advantage for solo learning. In help learning, there is no arrangement data and likewise, the expert can acquire from their own knowledge. Honestly, it gathers the arrangement models by experimentation while it is attempting its endeavour.

## 2 Related Literature

Agostino Sturaro proposed a sensible model for disappointment spread in associated force and interchanges organizations, Clue, thinking about heterogeneity of organization components, complex coherent and actual interdependencies, just as intricate disappointment proliferation conditions. They embraced manufactured and genuine organizations and thought about arbitrary and focused on assaults. Disappointment location and re-enactment for just digital actual organizations and framework frameworks [1].

Mohammadi Rouzbahani [2] concentrated peculiarity location which is a significant information examination task as one of the methodologies for CPSs security. Therefore, extraordinary inconsistency location strategies are introduced; however, it is hard to analyse the benefits and detriments of these methods. At long last, in this

AI (ML) strategies for the location of peculiarities are introduced through a contextual investigation that exhibits the viability of AI procedures at characterizing Bogus Information Infusion (FDI) assaults.

R. J. LA noticed the effect of inconstancy and relationships in levels of specialists on the power of reliant frameworks on the likelihood of falling disappointments with the 2003 power outage in Italy as a genuine model. The two of them can impact the probability of having calamitous disappointments in complex frameworks including different heterogeneous frameworks through reliance among the specialist. Specifically, these outcomes uncovered that both expanding changeability and positive reliance render the framework more powerful against arbitrary disappointment. Their discoveries have constraints on consolidating other diagram properties showed by both normal and designed frameworks, like assortative and bunching, and to comprehend their job in the vigour of reliant frameworks [3]. Dan Zhang introduced a study on ongoing advances on security issues of mechanical digital actual frameworks (ICPSs). They explicitly talked about two normal sorts of assaults, i.e. Forswearing of-Administration (DoS) assault and Duplicity assault, and present ongoing outcomes as far as assault location, assessment, and control of ICPSs. Groupings of current examinations are investigated and summed up dependent on various framework displaying and investigation techniques. Moreover, the benefits and hindrances of different strategies are likewise examined. At last, this paper finishes up with some potential future exploration bearings on secure ICPSs. The investigation is done distinctly on a nearby examination around impartial reliance charts and doesn't work past the neighbourhood of unbiased diagrams [4].

Domenico Cotroneo proposed a method for examining execution hints of dispersed frameworks under flaw infusion, by contrasting the executions with deficiency free ones to call attention to irregularities. To address the issue of non-determinism. The methodology with apparatuses for short coming infusion and troubleshooting, for example, for revealing the peculiarities to the clients and for bunching issue infusion tests to more readily uphold the human investigators. The disadvantage to the framework is less thought to Dangers Move Between the CPS and Clients which prompts numerous oddities [5].

Soumyadeep Thakur proposed a model that removes valuable highlights from the given highlights and afterwards utilizes a profound learning calculation to group the interruptions. An exceptional Nonexclusive Explicit autoencoder engineering where the conventional one learns the highlights that are normal across all types of organization interruptions and the particular ones learn highlights that are relating just to that space. The model has been assessed on the CICIDS2017 dataset, which is the biggest dataset of this sort accessible on the web, and we have set new benchmark results on this dataset. The model utilized in this space explicit to control matrix frameworks and future works is displaying it to every single sensible organization [6].

Dhir and Y. Kumar proposed research study is performed on the various structures utilized for Digital assault location utilizing the learning approach, which demonstrates the significance of AI and profound learning in Digital actual frameworks for distinguishing the dangers in a superior manner. The proposed concentrate on

various assaults has likewise featured the specialists to be more mindful of remarkable assaults that can turn out to be exceptionally perilous. Further, the investigation of different works done in breaking down various assaults is finished utilizing different methodologies and dataset that is covered alongside upsides and downsides to help in picking the best methodology as per the prerequisite [7].

CPS was recognized as one of the eight assessment need domains in August 2007 report of the President's Chamber of Counsellors on Science and Innovation, as CPS will be the middle portion of various essential structures and mechanical control systems soon. S. Han thought how to fittingly apply the interference revelation segment to CPS in this paper. By examining the surprising properties of CPS, they wanted to portray the specific necessities and a short time later, they arranged system of the interference disclosure part in CPS is introduced in regard to the layers of structure and express area techniques. Finally, some basic investigation issues are perceived for illuminating the following assessments [8].

This methodology by Peng et al. [9], shows that there exists an edge for the extent of defective hubs, past which the digital actual frameworks breakdown. It likewise decides the basic qualities for various framework boundaries. The dependability investigation done here dependent on network permeation hypothesis can be adequately used for hostile to assault and security purposes in coupled digital actual frameworks. However, this methodology neglects to decide the specific limit.

### 3 Proposed Method

The proposed system is a model which uses machine learning algorithms to detect the attacks on the given network data. The system analyses the existing dataset and processes it. It then using various feature selection and PCA methods provides a model for the detection of the vulnerabilities. When the network data is provided into the model, it detects any failures and classifies the attacks on basis of traffic congestion or other various attacks possible. The dataset is a combination of data generated using virtual network attack simulation and those collected from online data sources.

In this segment, we initially depict our dataset, its substance, and our way to deal with setting it up for the learning task. At that point, we disclose our way to deal with utilizing the information for preparing the model and the framework plan.

#### 3.1 Data Collection

At first to accumulate information, we have mimicked a virtual organization situation that makes failures and assaults utilizing Mininet in python. At that point, we have utilized telegraf and influxdb. In this telegraf is a measurements specialist answerable for gathering information about the host it is running on. It is altogether

module-driven. The last is an administration Framework whose engineering is explicitly outfitted towards measurements. The interconnection between the two is basic, together of telegraf's modules offers local help for influxdb. We are watching ICMP traffic to outline estimates about the state of the organization. We ran telegraf on an organization switch that was clearly connected with the controller where our influxdb model is running controller through HTTP. This suggests that the filesystem is shared across the association segments we make with Mininet and consequently the host machine itself. This host machine has direct accessibility with the VM working with the controller. We are having the chance to run a telegraf occurrence on Mininet's Host 4 whose input module will assemble ICMP information and whose yield will be an enter the VM's home index. This information is shipped off influxdb through the telegraf module by utilizing the python content.

The data gathered above was from virtual simulation and theses were combined with the data collected from the network APA-DDoS dataset from Kaggle [10] for making our model more general to various forms of failures and attacks. This dataset consists of data related to ACK and PUSH-ACK flooding DDoS attacks in network.

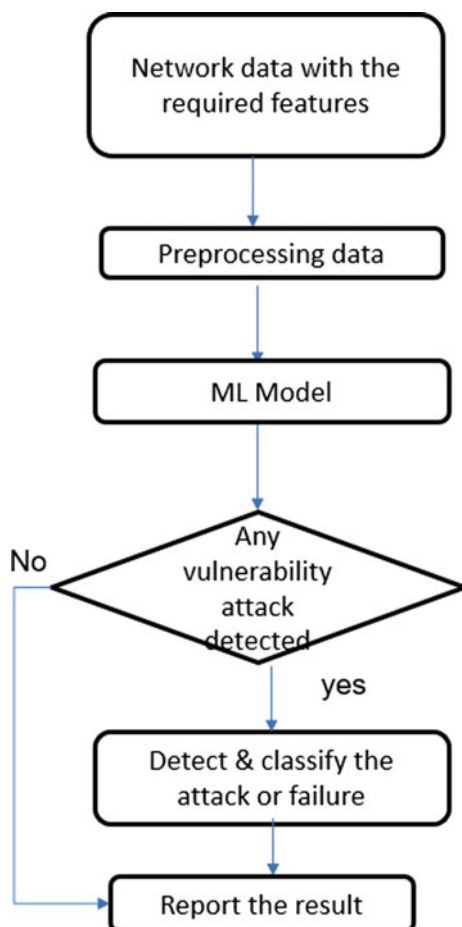
### 3.2 *System Architecture*

The diagram below (Fig. 1) depicts the overall comprehensive architecture design of our proposed method's framework.

The working of the total model from the above diagram can be explained as follows. The data is collected from virtually simulated network and network traffic dataset in Kaggle are pre-processed. The data is again trained to testing vulnerabilities and then data will be checked with the ML model. Now if any network simulation data is given into the model it checks for any anomalies it detects and then classifies the type of attack based on the data.

### 3.3 *Data Pre-processing*

Data pre-processing is a cycle of setting up the crude information and making it appropriate for an AI model. It is the first and significant advance while making a machine learning model. Data by and large contains clamours, missing qualities, and possibly in an unusable organization that can't be straightforwardly utilized for machine learning models. Data pre-processing is a necessary assignment for cleaning the data and making it reasonable for a machine learning model which likewise expands the exactness and accuracy of an machine learning model. By utilizing the different pre-processors and autoencoders, the data utilized is eliminated from invalid information and oddities in it. This serves to at first train the model precisely and furthermore this interaction is done on the approaching contribution from CPS Framework to the model to meet the model element necessities.

**Fig. 1** System flow diagram

### ***3.4 Machine Learning Module with Random Forest Classifier***

The data is sent to the model after it has been pre-processed, built using the RFC for the attacks and failure detection. The dataset basically consists of various columns of data such as ipsrc, ipdest, port numbers, frames, Tx packets, bytes, etc. This is first scaled using the min-max scaler to process using the machine learning algorithm. Minmax Scaling is done to do normalization in the data.

### 3.5 *The Random Forest Classifier*

Random forest, similar to its name suggests, comprises of an outsized number of individual choice trees that work as an outfit. Every individual tree inside the random forest lets out a classification forecast and in this manner the class with the principal votes turns into our model's expectation. The major idea driving random forest might be a straightforward yet incredible one—the astuteness of groups. In information science-talk, the explanation that the arbitrary timberland model functions admirably is that an enormous number of moderately uncorrelated models (trees) working as a council will beat any of the individual constituent models.

In the calculation, 75% of the preparation dataset is utilized for preloading and this is utilized in tree advancement. The leftover preparing dataset is utilized to test for blunders. The branch to be produced from every hub is controlled by the GINI list esteem utilized in the CART algorithm which is utilized for the tree advancement measure from the acquired datasets in a random forest classifier.

This algorithm is more useful in the classification problems to handle large normalized data. This is the reason why we choose this model. It helps us best to detect and classify the various forms of attacks.

## 4 Execution Results and Discussions

In this section, we assess the capability of a model based on classification algorithm on the readied dataset and its results. The experiments are implemented on a Windows 64-bit system with a minimum of 4 GB RAM and 1.25GHZ processor frequency. Python 3.7 is used to code all scripts for extracting and pre-processing data, as well as training the model. Also for experimenting purposes, Principal component analysis done on the dataset and then normalized and trained. This model is used to deploy in the flask app for demonstration purposes.

For evaluating the model trained with random forest classifier, we use the confusion matrix. A confusion matrix is a table that depicts the performance of a classification model on a set of test data for which the true characteristics are identified. The values in the confusion matrix are used to measure evaluation metrics. The values in the confusion matrix are summarized as follows:

**True-positive (TrPo):** The number of data instances that are labelled as yes and classified as yes correctly in the prediction class.

**True-negative (TrNe):** The number of data instances that are named or labelled as no in actual and classified as no correctly in prediction class.

**False-negative (FaNe):** The number of data instances that are named or labelled as no and classified as yes incorrectly in prediction class.

**False-positive (FaPo):** The number of data instances that are named or labelled as yes and classified as no incorrectly in prediction class.

**Table 1** Execution Metrics of the Classifier Model

Classifier	Accuracy	Precision	Recall	F-Score
Random forest (%)	96	95.21	94.71	94.89

Following the acquisition of the confusion matrix results, the following is a summary of the assessment models developed using those qualities.

**Accuracy** The proportion of the number of tests correctly defined by the system to the total number of tests is known as accuracy esteem. It is determined as follows,

$$\text{Accuracy} = (\text{TrPo} + \text{TrNe}) / (\text{TrPo} + \text{FaPo} + \text{FaNe} + \text{TrNe}) \quad (1)$$

**Precision (P)** Precision is defined as the ratio of correctly predicted positive perceptions to all predicted positive perceptions.

$$\text{Precision} = \text{TrPo} / (\text{TrPo} + \text{FaPo}) \quad (2)$$

**Recall (Sensitivity)** Recall is characterized as the ratio of effectively anticipated positive outcomes to all perceptions in a genuine class -yes.

$$\text{Recall} = \text{TrPo} / (\text{TrPo} + \text{FaNe}) \quad (3)$$

**F1 score** The weighted normal of Precision and Recall is the F1 Score.

$$\text{F1 Score} = 2 * (\text{Sensitivity} * P) / (\text{Sensitivity} + P) \quad (4)$$

Table 1 summarizes the results of the experiments. We used Scikit—Learn for the implementation of the above models and metrics. We could see that the model takes in the input and gives out the prediction result as which type of attack has taken place as benign or DDoS-ACK attack, etc.

## 5 Conclusion and Future Works

In the interdependent cyber-physical structures, we suggested a practical model for intrusion. Our model, takes in the network data with various features of it and predicts for any vulnerability and detects it. In case, it classifies the attack based on the data with the trained model. Security is quite possibly the main factors in CPS systems on account of the recurrence of revealed digital assaults. Thus, detection methods were presented. Finally, showing the effectiveness of the ML algorithm in classifying cyber-physical systems attack was given. Our result also demonstrated that feature selection might provide a way for further deep analysis.



We used random forest classifier to classify the attacks. Also a wide variety of dataset can be found in real time and used to get better rational results. In further, these can be demonstrated with the help of precise feature selection and neural networks for better results. Also some real time simulated frameworks can be developed to avoid cascading effects in CPS systems with failure propagation detection and avoidance. This can be integrated with the logs of a system to generate alert and report the defaults.

## References

1. Sturaro A, Silvestri S, Conti M, Das SK, A realistic model for failure propagation in interdependent cyber-physical systems, IEEE. <https://doi.org/10.1109/TNSE.2018.2872034>
2. Mohammadi Rouzbahani H, Karimipour H, Rahimnejad A, Dehghantanha A, Srivastava G. (2020) Anomaly detection in cyber-physical systems using machine learning. In: Choo KK, Dehghantanha A. (eds) Handbook of big data privacy. Springer, Cham
3. La RJ (2017) Cascading failures in interdependent systems: Impact of degree variability and dependence. IEEE Trans Netw Sci Eng
4. Zhang D, Wang QG, Feng G, Shi Y, Vasilakos AV (2021) A survey on attack detection, estimation and control of industrial cyber–physical systems. ISA Trans
5. Cotroneo D, Simone LD, Liguori P, Natella R (2019) Enhancing failure propagation analysis in cloud computing systems. In: 2019 IEEE 30th international symposium on software reliability engineering
6. Thakur S, Chakraborty A, De R, Kumar N, Sarkar R (2021) Intrusion detection in cyber-physical systems using a generic and domain specific deep autoencoder model. Comput Elect Eng 91
7. Dhir, Kumar Y (2020) Study of machine and deep learning classifications in cyber physical system. In: 2020 third international conference on smart systems and inventive technology (ICSSIT). Tirunelveli, India
8. Xie HM, Chen H, Ling Y (2014) Intrusion detection in cyber-physical systems: techniques and challenges. IEEE Syst J 8(4):1052–1062
9. Peng H, Kan Z, Zhao D, Han J (2019) Security assessment for interdependent heterogeneous cyber physical systems. Mobile Netw Appl
10. <https://www.kaggle.com/yashwanthkumbam/apaddos-dataset>

# Comparative Study to Perform Tweets/NEWS Classification Using Big Data Analytics for Predicting Stock Market Price Movement



Dhara N. Darji, Satyen M. Parikh, and Hiral R. Patel

**Abstract** In finding prediction in a range of domains, sentiment classification of the Twitter and NEWS data has been applied successfully. However, it is still difficult and continuing research to use sentiment classification to forecast stock market factors. Sentiment analysis is complex and essential task that must be done, and natural language processing and machine learning all are concerned. The purpose of this research is to compare accuracy of the three feature extraction techniques (TF-IDF, CountVectorizer, Word2Vec) and three machine learning classifiers (logistic regression, random forest, and Naive Bayes classifier) in order to provide stock-related tweets and news data with positive, negative, and neutral sentiment. The experimented results show that random forest algorithm provides highest accuracy and logistic regression algorithm with the less accuracy.

**Keywords** Spark NLP pipeline · Preprocessing · Tokenizer · Feature extraction · Machine learning classifier · Sentiment classification

## 1 Introduction

Due to the exponential growth of Web and smartphone applications, a massive amount of data is generated every second. The people use social media platform to share their feelings or opinion with friends. These data may represent an extremely important source of information for many fields like merchants, stock market, and in election works among them the stock market is a fascinating area to research. Every second, the stock market and social media change. Twitter's patterns will have an effect on

---

D. N. Darji (✉) · H. R. Patel

Department of Computer Science, Ganpat University, Gujarat, India

e-mail: [dnd01@ganpatuniversity.ac.in](mailto:dnd01@ganpatuniversity.ac.in)

H. R. Patel

e-mail: [hrp02@ganpatuniversity.ac.in](mailto:hrp02@ganpatuniversity.ac.in)

S. M. Parikh

Faculty of Computer Applications, Ganpat University, Gujarat, India

e-mail: [satyen.parikh@ganpatuniversity.ac.in](mailto:satyen.parikh@ganpatuniversity.ac.in)

the stock market. We will find the correlation between stock price and tweets & daily news. But the tweets and news data are available in raw text data and machine learning algorithms are unable to comprehend this raw data. Preprocessing of the data is needed. Natural language processing (NLP) is used in the fundamental research approach to interpret tweets and financial news details to find positive, negative, and neutral sentiments depending on the contents of the texts.

In this paper, we created multiple sentiment classification models for sentiment analysis. An analytical comparison of three function extraction methods is presented (TF-IDF, CountVectorizer, and Word2Vec). Using three machine learning classifiers (logistic regression, random forest, and Naive Bayes (NB)), the classification accuracy of the feature extraction techniques is analyzed. Finally, the developed models are tested against an opinion corpus of Twitter tweets and news gathered from a common Web site about financial or stock market data. The model is built using the Apache Spark big data framework. The Apache Spark group developed PySpark to help Python users work with Spark. To build a model Spark NLP text processing library used for advanced natural language processing, while MLlib is a collection of common machine learning algorithms used for classification.

The rest of this paper is laid out as follows. Section 2 of the paper examines related work in the area of sentiment classification. In Sect. 3, define the proposed model's architecture. Section 4 describes the applicable methods for constructing the Spark NLP pipeline, providing a discussion of the different steps as well as model assessment. The comparative study of experimental results is presented in Sect. 5. Finally, in Sect. 6, we conclude our work and explore future research directions.

## 2 Literature Survey

There are several detailed analyses of sentiment analysis in the research literature that define the various approaches used for sentiment analysis in textual data. In Naresh and Venkata Krishna [1], an optimization-based machine learning algorithm has been used for classification of Twitter data by applying three stages, data preprocessing, feature extracting, and classification by applying machine learning algorithm, respectively.

In Al-Amrani et al. [2], author compared five machine learning algorithms for classifying the SMS text messaging dataset and concluded that logistic regression algorithm gives the better accuracy rather than other four algorithms.

Devika et al. [3] show the costs and benefits of the rule-based, lexicon-based, and machine learning approaches for sentiment classification as well as also discuss the different levels of analyzing sentiments including semantics by using n-gram evaluation instead of word-by-word analysis.

In Tripathy et al. [4], author used textual movie reviews and converts it into numerical matrix to identify the features for testing and training the model and discusses the confusion matrix, and various performance assessment factors retrieved using Naive Bayes and support vector machine algorithm also conclude that in future

the other similar classification strategies under supervised learning methodology are implemented and result with the SVM is compared.

People are constantly using emoticons in their text to show their emotions. In Ullah et al. [5], implement the sentiment analysis on Twitter-based airline data with only text and text and emoticon both. Various features like TF-IDF, bag of words, N-gram, and emoticon lexicons and machine learning and deep learning algorithms applied on that determine the sentiments and finally conclude that use of emoticons in conjunction with text and applying DL algorithm seems with better accuracy.

The effects of various feature selection and machine learning approaches in terms of the macro-F measure for Arabic sentiment classification were investigated in Omar et al. [6], and it was determined that SVM-based FSM performs the best with an accuracy of 92.4%, while the NB classifier works best to be used with the other FSMs evaluated.

In Rane and Kumar [7], the airline dataset was classified into three sentiment class using seven machine learning techniques: positive, negative, and neutral. To execute deep learning algorithms, the input data was preprocessed and the tweets were interpreted as vectors. AdaBoost has decent precision of 84.5%, according to observation. The biggest disadvantage being that only a small number of tweets is considered but if the number of tweets is raised, high precision will be possible.

### 3 Proposed Model

The propose model contains the various stages for sentiment classification of the stock market textual data (Fig. 1).

A single execution Spark NLP pipeline for our model.  
pipeline = Pipeline(

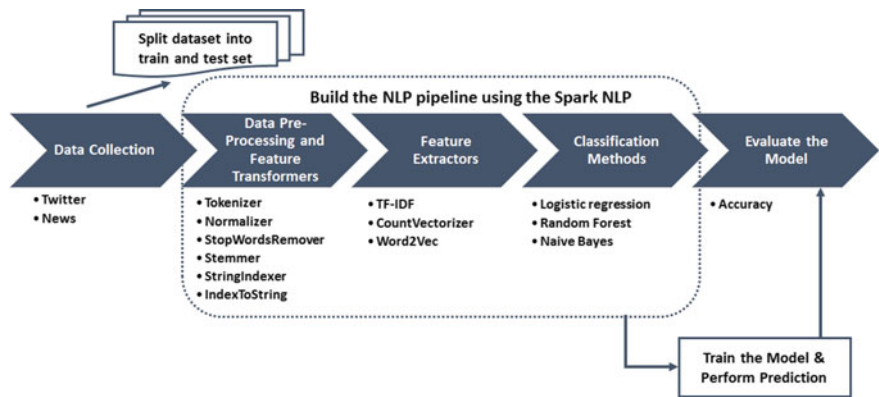


Fig. 1 Proposed model

```

stages = [document_assembler,
          tokenizer,
          normalizer,
          stopwords_cleaner,
          stemmer,
          finisher,
          hashingTF,
          idf,
          countVectorizer,
          word2vec
          label_stringIdx,
          classifier,
          label_to_stringIdx])

```

## 4 Applicable Methodology

### 4.1 Data Collections

For sentiment classification, the input data is collected from the Twitter and news data sources. Python script is created using Tweepy API for Twitter data, and the appropriate hashtag passes as parameters for collecting companywise stock or financial tweets. The Moneycontrol and CNBC Web sites various rss feed [Top News, Latest News, Result News, Economy News] is used for the financial news data [4].

### 4.2 Data Preprocessing and Feature Transformers

Preprocessing is the process of cleaning the data by extracting noisy, redundant, and unreliable data from a dataset and prepare the text for the better classification. The gathered tweets include a variety of emotions, hash tags, and special symbols, URL and RT [retweets] as well as a comment. Remove all mention noisy data from the tweet dataset because they are considered distracting and meaningless [5].

Features transformation is a process of scaling, converting features from one representation to another, or modifying features.

There are several feature transformers that are used in Spark NLP for preprocessing data:

**Tokenizer.** Tokenization is the act of breaking down a piece of text (sentence) into individual terms or words. This functionality is given by a basic Tokenizer class. The Tokenizer works with DocumentAssembler or SentenceDetector output because it requires a Document annotation type. The first text column of the Spark data frame (df) is fed into the DocumentAssembler() transformer, which creates

a new Document type column called “document.” This is the initial entry point for every Spark data frame into Spark NLP. The text column is then passed to `Tokenizer()`, which tokenizes each sentence and creates a new Token type column called “token.”

**Normalizer.** Normalizer transformer used to translate a dataset of vector rows into a unit norm by normalizing each Vector. The normalizer adopted the regex template and transformed the words based on the given directory to delete all dirty characters from the document.

**StopWordsRemover.** Stop words are words that can be omitted from the input because they occur often and have no meaning. `StopWordsRemover` takes a series of strings as input (In this paper, the output of a `Tokenizer` takes as an input) and eliminates all stop words.

**Stemmer.** The Stemmer Annotator is used to split out a portion of a word or reduce it to its stem or base.

**StringIndexer.** `StringIndexer` transforms a string column of labels to a column of label indices. String indexing is a very easy idea to understand. We basically assign a number to each category and use this number in our models instead of the label.

**IndexToString.** `IndexToString` converts a column of label indices back to a column with the initial labels as string. To train the model with indices, we typically use the `StringIndexer` to create indices from labels, and then we use `IndexToString` to retrieve the original labels from the column of forecast indices.

### 4.3 Feature Extractors Techniques

In the feature extraction phase, text-based tweet and news data is transformed into numerical data in the form of a feature vector. The latter is a machine learning technique applied on these numerical features.

In this paper, the following methodologies were used to vectorize textual data into numerical vectors.

**TF-IDF.** TF-IDF factor indicates how important a term to a document in the corpus. The definitions term frequency and document frequency are described in a variety of ways. In `MLlib`, `HashingTF` Transformer and `IDF Estimator` are used, respectively, to perform the TF and IDF. Separating TF and IDF makes them more flexible. `HashingTF` is used to convert sets of terms into a fixed-length feature vectors and IDF scales each feature based on the feature vectors created by `HashingTF`. Intuitively, it down-weights features which appear frequently in a corpus.

**CountVectorizer.** `CountVectorizer` assists in the translation of a set of text documents into token count vectors.

**Word2Vec.** The `Word2Vec` extractor converts each word into a vector. After that, the vector can be used to make predictions.

4.4 Classification Methods

In this paper, three classifier methods logistic regression, random forest, and Naive Bayes are used for sentiment classification of stock market textual data collected from the Twitter and news.

4.5 Evaluate the Model

The several experiments are conducted to evaluate our model. To evaluate the performance or accuracy of the three classification methods according to the three features extraction techniques, Spark NLP by John Snow Labs has a number of annotators for performing specialized NLP activities. Most machine learning models are available in Apache Spark MLlib, as well as a variety of feature transformers such as Tokenizer and StopWordRemover as well as feature extractors such as CountVectorizer, TF-IDF, and Word2Vec. While these transformers and extractors are necessary to construct a simple NLP pipeline, more sophisticated techniques such as stemming, lemmatization, normalization are needed to build a more detailed pipeline so we can say that the big data framework Apache Spark with MLlib used to build the Spark NLP Pipeline, which developed with single execution including all major three stages of model like data preprocessing and transformation, feature extraction, and classification. Finally based on the accuracy factor, evaluate the model (Figs. 2 and 3, Table 1).

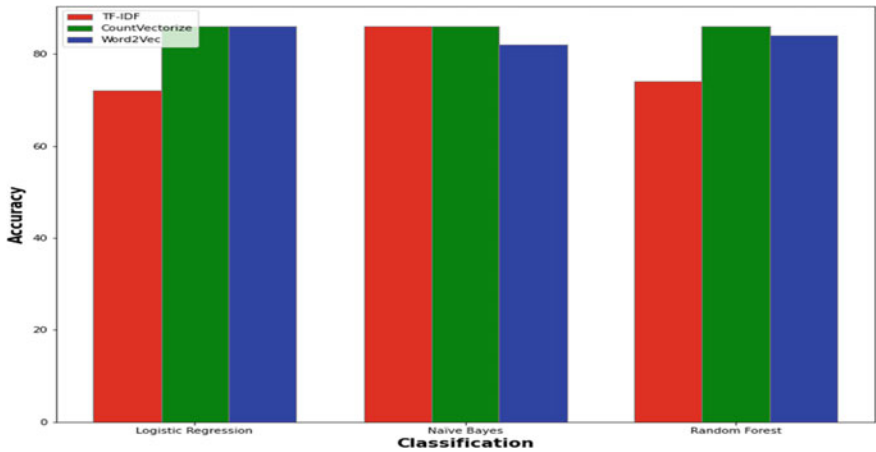
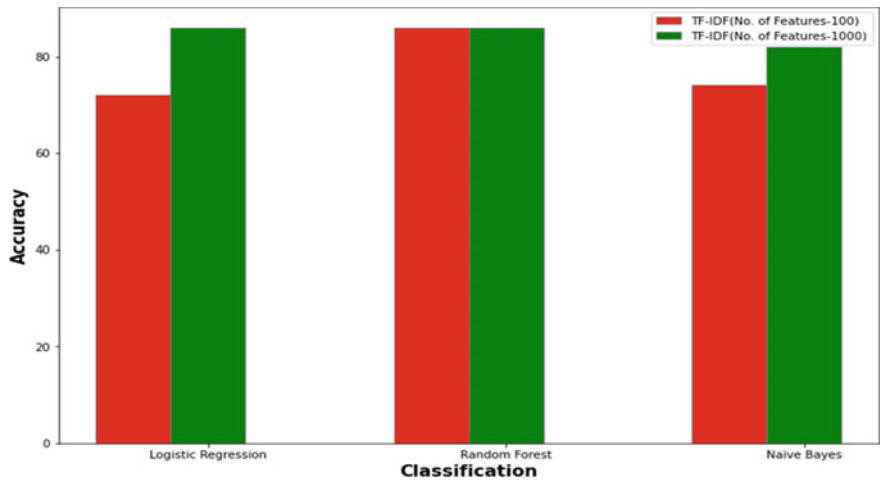


Fig. 2 Performance comparison for classification



**Fig. 3** According to the no. of features performance comparison for classification using TF-IDF technique

**Table 1** Comparative analysis of sentiment classification using three classifiers and three feature extraction techniques with respect to their accuracy

Classifier	Feature extraction	Accuracy (%)
Logistic regression	TF-IDF (100 features)	72
	TF-IDF (1000 features)	86
	CountVectorizer	86
	Word2Vec	86
Random forest	TF-IDF(100 features)	86
	TF-IDF(1000 features)	86
	CountVectorizer	86
	Word2Vec	82
Naive Bayes	TF-IDF(100 features)	74
	TF-IDF(1000 features)	82
	CountVectorizer	86
	Word2Vec	84

4.6 Conclusion and Future Direction

This paper provided a comprehensive comparison of three feature extraction approaches and three machine learning algorithms for sentiment classification. The analysis of the results of various feature extraction and machine learning approaches in terms of accuracy is the key contribution of this work. The final result indicates that if feature extract through CountVectorizer technique gives same accuracy for all the classifier, word2vec performs best with the logistic regression classifier. Lastly the TF-IDF seems better with random forest and it gives the same accuracy with different



no. of features. Only three different classifiers were used in this research. Other related classification methods based on supervised learning and neural networks could be considered in the future.

## References

1. Naresh A, Venkata Krishna P (2020) An efficient approach for sentiment analysis using machine learning algorithm. *Evol Intel*. <https://doi.org/10.1007/s12065-020-00429-1>
2. Al-Amrani Y, Lazaar M, Elkadiri KE (2017) Sentiment analysis using supervised classification algorithms. In: *Proceedings of the 2nd international conference on big data, cloud and applications*. <https://doi.org/10.1145/3090354.3090417>
3. Devika M, Sunitha C, Ganesh A (2016) Sentiment analysis: a comparative study on different approaches. *Procedia Comput Sci* 87:44–49. <https://doi.org/10.1016/j.procs.2016.05.124>
4. Tripathy A, Agrawal A, Rath SK (2015) Classification of sentimental reviews using machine learning techniques. *Procedia Comput Sci* 57:821–829. <https://doi.org/10.1016/j.procs.2015.07.523>
5. Ullah MA, Marium SM, Begum SA, Dipa NS (2020) An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express* 6(4):357–360. <https://doi.org/10.1016/j.ict.2020.07.003>
6. Omar N, Albared M, Al-Moslmi T, Al-Shabi A (2014) A comparative study of feature selection and machine learning algorithms for arabic sentiment classification. *Inf Retrieval Technol*, pp 429–443. [https://doi.org/10.1007/978-3-319-12844-3\\_37](https://doi.org/10.1007/978-3-319-12844-3_37)
7. Rane A, Kumar A (2018) Sentiment classification system of Twitter data for US airline service Analysis. In: *2018 IEEE 42nd annual computer software and applications conference (COMPSAC)*. <https://doi.org/10.1109/compsac.2018.00114>
8. Bhavitha BK, Rodrigues AP, Chiplunkar NN (2017) Comparative study of machine learning techniques in sentimental analysis. In: *2017 International conference on inventive communication and computational technologies (ICICCT)*. <https://doi.org/10.1109/icicct.2017.7975191>
9. Pai P, Liu C (2018) Predicting vehicle sales by sentiment analysis of twitter data and stock market values. *IEEE Access* 6:57655–57662. <https://doi.org/10.1109/access.2018.2873730>
10. Wang L, Niu J, Yu S (2020) SentiDiff: Combining textual information and Sentiment DIFFUSION patterns for Twitter sentiment analysis. *IEEE Trans Knowl Data Eng* 32(10):2026–2039. <https://doi.org/10.1109/tkde.2019.2913641>
11. Das S, Behera RK, Kumar M, Rath SK (2018) Real-time sentiment analysis of twitter streaming data for stock prediction. *Procedia Comput Sci* 132:956–964. <https://doi.org/10.1016/j.procs.2018.05.111>
12. Derakhshan A, Beigy H (2019) Sentiment analysis on stock social media for stock price movement prediction. *Eng Appl Artif Intell* 85:569–578. <https://doi.org/10.1016/j.engappai.2019.07.002>
13. Nguyen TH, Shirai K, Velcin J (2015) Sentiment analysis on social media for stock movement prediction. *Expert Syst Appl* 42(24):9603–9611. <https://doi.org/10.1016/j.eswa.2015.07.052>
14. Khan W, Ghazanfar MA, Azam MA, Karami A, Alyoubi KH, Alfakeeh AS (2020) Stock market prediction using machine learning classifiers and social media, news. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-020-01839-w>
15. Pathak A, Shetty NP (2018) Indian stock market prediction using machine learning and sentiment analysis. *Adv Intell Syst Comput*, pp 595–603. [https://doi.org/10.1007/978-981-10-8055-5\\_53](https://doi.org/10.1007/978-981-10-8055-5_53)

16. Romanowski A, Skuza M (2016) Towards predicting stock price moves with aid of sentiment analysis of Twitter social network data and big data processing environment. *Advances in Business ICT: New Ideas from Ongoing Research*, pp 105–123. [https://doi.org/10.1007/978-3-319-47208-9\\_7](https://doi.org/10.1007/978-3-319-47208-9_7)

# Design and Improvement in the Performance of Multi-band House-Shaped Microstrip Patch Antenna with SRR-Based Metamaterial for 5G Applications



John Colaco and Rajesh B. Lohani

**Abstract** Microstrip patch antennas embrace a major role in 5G wireless communication technology because of their ease in manufacturing, high reliability, and high-power efficiency. The incorporation of metamaterials will help to boost the efficiency of microstrip patch antennas by improving their different characteristics such as S-parameter, bandwidth, gain, and power density. In this regard, using FEKO software, the authors have first designed a multi-band house-shaped microstrip patch antenna at 27.5 GHz, and then designed split-ring resonator-based metamaterial and integrated said SRR-based metamaterial onto the substrate and on the patch for investigating the performance of the proposed microstrip antenna. During the investigation, it is found and observed that the bandwidth is enhanced from 4.769 to 5.370 GHz, return loss is enhanced from  $-19.42$  to  $-21.79$  dB, the gain is enhanced from 5.36 to 5.6 dBi, and power density is enhanced from 57.66 to 66.17 dBW/m<sup>2</sup>. In this design, the authors proposed an RT/Duroid substrate of 5880 with a relative permittivity of 2.2 and a tangent dielectric loss of 0.014.

**Keywords** Microstrip patch antenna · 5G · SRR (Split-ring resonator) · Metamaterial · Negative permittivity and permeability · Absorptivity · RCS reduction · Bandwidth · Return loss · Gain · Power density

## 1 Introduction

The findings of a novel, man-made artificial material like metamaterial have revolutionized the electronics and wireless communication system because of its extraordinary characteristics like negative electrical (permittivity) and magnetic (permeability) properties causing the electromagnetic wave to propagate in a backward direction in the DNG medium which is generally not observed in any other materials

---

J. Colaco (✉) · R. B. Lohani

Department of Electronics and Telecommunication Engineering, Goa Engineering College, Ponda, Goa, India

R. B. Lohani

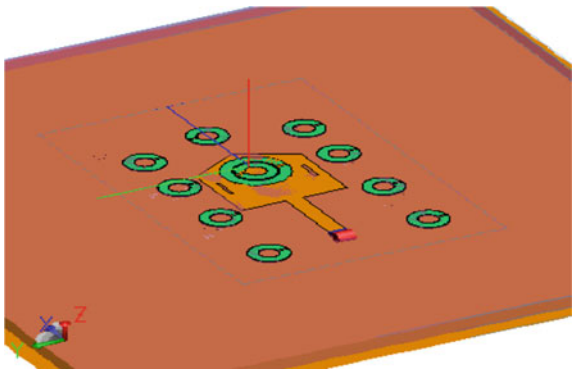
e-mail: [rblohani@gec.ac.in](mailto:rblohani@gec.ac.in)

of nature. In this medium, a thin-wire-based ENG (epsilon-negative) material and split-ring resonator (SRR)-based MNG (mu-negative) materials are both combined to satisfy both properties of magnetic ( $\mu < 0$ ) and electric material left ( $\epsilon < 0$ ) in the third quadrant [1]. The impacts produced by the metamaterials can be observed amid the transmission of electromagnetic waves such as wave propagation [1]. This phenomenon transpires due to an antiparallel group, and the phase velocity leads to the inversion of the wavefronts, while its energy is moving away from the source. The characterization of metamaterial is done by utilizing Maxwell equations. These Maxwell equations give the relationship between electric and magnetic fields that helps in determining this negative permittivity and negative permeability, respectively, when interacted with a double-negative medium which is also known as negative refractive index [1]. This artificial material has shown tremendous support in the design of microstrip patch antennas by improving its characteristics, thus further extending its support in the 5G wireless communication system. The microstrip patch antennas are low-cost, easily manufactured, fabricated, and compact size antennas. Consequently, the request for a microstrip patch antenna for 5G applications is rising exceptionally quick. The main inspiration for using metamaterials in microstrip patch antennas for 5G communication has been the essential of large bandwidth and high-energy efficiency. This 5G is the inventive high-speed and highly reliable fifth-generation remote communication way better than the existing 4G. The major 5G applications include media and virtual experience and massive connectivity machine-type communication which includes online learning and teaching, e-health, smart cities [2]. The metamaterial used in the proposed design is a split-ring resonator which is a unit cell composite microstructure that produces unique electromagnetic properties such as negative permittivity and permeability which is generally not found in nature. This SRR is composed of two pair of concentric rings made of copper with a gap between them. The concentric ring represents the inductor ( $L$ ), and the gap represents the capacitor ( $C$ ). Hence, this structure forms an artificial magnetic LC resonator which determines magnetic polarization when current flows through it. This structure when incorporated on the microstrip patch shows improvement in the characteristics such as bandwidth, return loss, beam-width (gain), and power radiations. The purpose of this proposed work is to design a unique microstrip patch antenna using Rogers RT Duroid substrate for 5G applications and to etch two vertical slots on the patch as these slots produce currents on the patch to produce multiple resonances and then improve the performance of proposed antenna by loading and incorporating this SRR metamaterial on this patch and substrate as made known in Fig. 1 which showed improvement in bandwidth, beam-width, gain, and power density.

## 2 Related Work

The authors have conducted an extensive survey with regard to 5G technology and its development [3]. The authors have designed a reconfigurable contiguous square

**Fig. 1** Proposed 3D view of microstrip house-shaped patch antenna



resonator-based metamaterial (MM) structure for enhancing the gain of the antenna at 28 GHz for 5G applications [4]. The authors have proposed a microstrip patch antenna with a complementary rhombus-shaped split-ring resonator for enhancing bandwidth [5]. The authors have proposed a multi-band rectangular microstrip patch antenna in the frequency band 1–10 GHz and incorporated with squared split-ring resonator metamaterial on the rear-side of antenna substrate to improve the execution of the antenna [6]. The authors have proposed a multi-band square-shaped microstrip patch antenna for 5G communications [7]. The authors have proposed a multi-band microstrip patch antenna at 2.4 GHz for IoT-based biomedical intelligent system in agriculture [8].

### 3 Methodology

#### 3.1 Design and Investigation of SRR-Based Metamaterial

Table 1 depicts the parameters’ values obtained for the design.

Figure 2 represents the proposed 3D view of split-ring resonator.

**Table 1** Design antenna parameters

S. No.	Details of parameters	Value (mm)
1	Gap between two concentric rings	0.2
2	Radius of the outer concentric ring	1.8
3	Radius of the inner concentric ring	1.5
4	Width of the concentric rings	0.4
5	Gap for each of the concentric ring	0.3
6	Height of the substrate	0.55
7	Length and breadth of substrate	6.2*6.2

**Fig. 2** Proposed 3D view of split-ring resonator

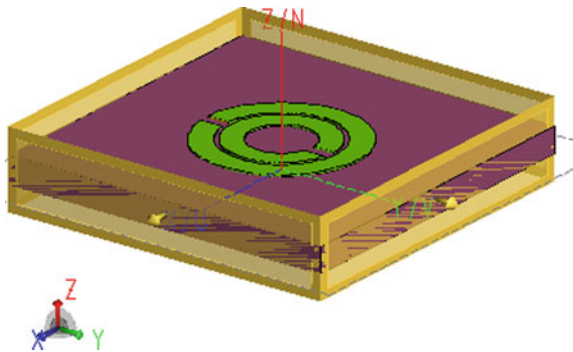
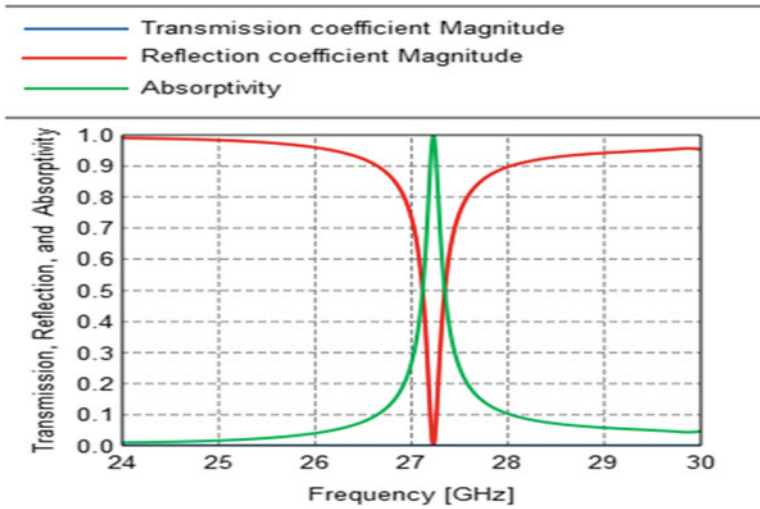


Figure 3 represents the absorptivity due to the magnitude of transmission and reflection coefficient when the plane electromagnetic wave is incident normally in the double-negative medium with proper electromagnetic boundary conditions, thus causing the scattering phenomenon of the plane wave and the refractive index become less than zero [9]. This absorptivity of SRR-based metamaterial will increase the radiating capability of the proposed microstrip antenna as well as sensitivity. The absorptivity observed is 100% and is obtained using [8].

Figure 4 represents the properties of SRR showing negative permittivity and negative permeability indicating the strength of SRR metamaterial causing polarization of the electric field and magnetization of the magnetic field in the frequency domain, respectively, which is unseen compared to other materials of nature. These properties help to improve the performance of the microstrip antenna [10].



**Fig. 3** Absorptivity

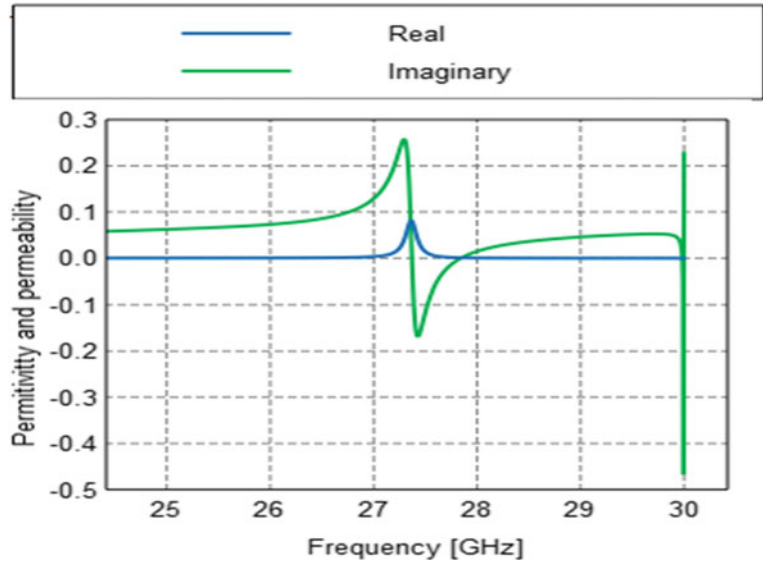


Fig. 4 Negative permittivity and permeability

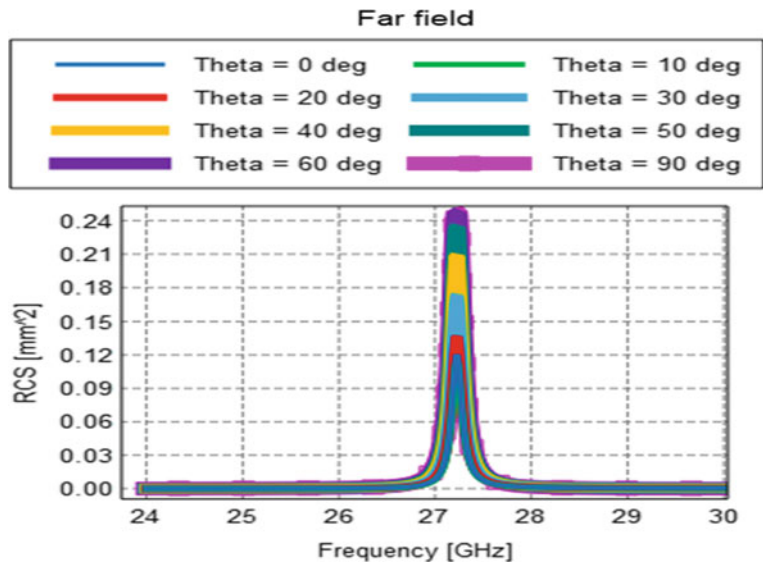
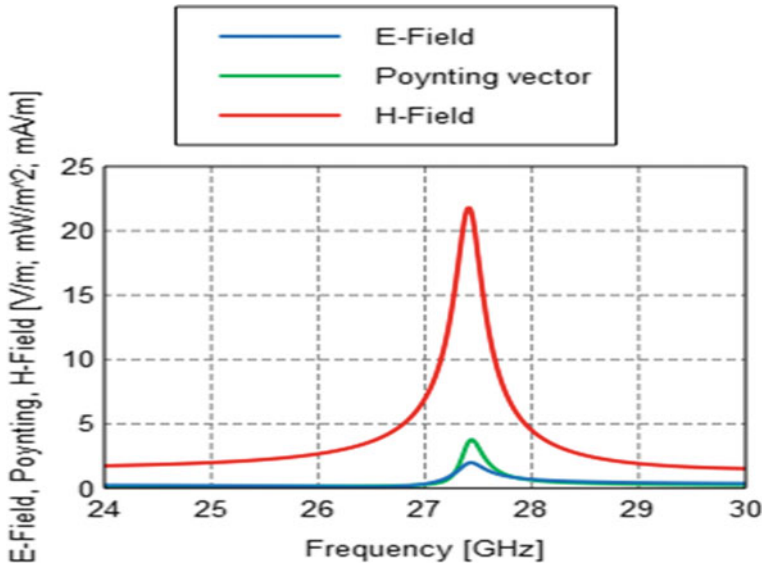


Fig. 5 RCS reduction with respect to theta



**Fig. 6** Power density

Figure 5 represents the radar cross section reduction at different angles of theta which signifies the SRR metamaterial as absorbing material that reduces the backward wave scattering through the transformation of the incident electromagnetic (EM) energy into heat [11].

Figure 6 represents power density or poynting vector which is due to the effect of electric and magnetic field intensity on SRR metamaterial that signifies a good efficient amount of energy radiates per unit area uniformly in all directions.

### ***3.2 Design and Investigation of Microstrip House-Shaped Patch Antenna with SRR-Based Metamaterial and Without SRR Metamaterial***

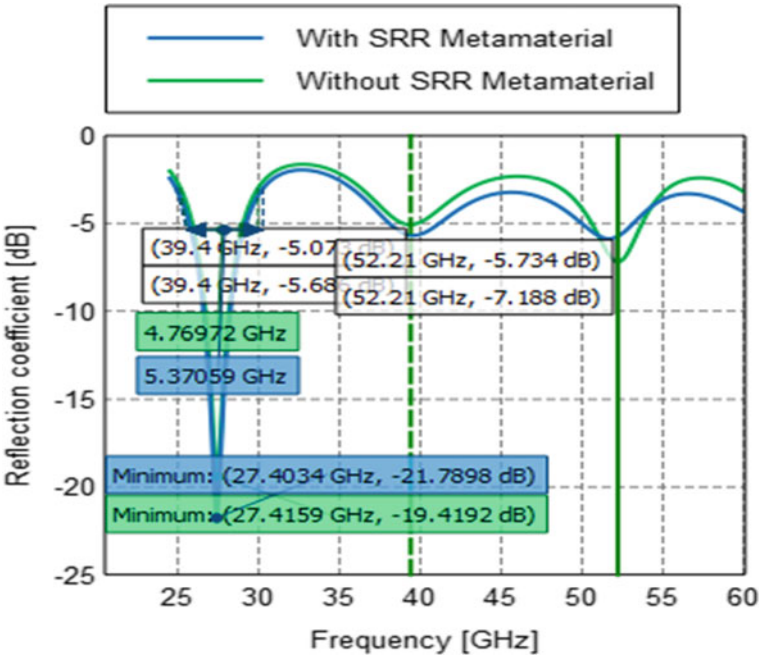
The proposed house-shaped microstrip patch antenna consists of a thin metallic patch placed on a dielectric substrate having a ground plane larger than the size of the patch and is feed with a 50-ohm microstrip transmission line to supply power to the antenna elements. Table 2 show the values of the parameters used in the proposed design.

Figure 7 shows the graph of return loss and bandwidth which are the most important characteristics of microstrip patch antenna where return loss indicates that the proposed antenna radiates more than 90% of the power supplied to the antenna to that of 10% or less power is reflected by the antenna measured in dB (decibel) and bandwidth indicates the rate at which the antenna radiates in 20–60 GHz of range of



**Table 2** Design antenna parameters

Parameter symbols	Details of parameters	Value (mm)
$W_s$	Breadth of the substrate	12
$L_s$	Length of the substrate	12
$W_f$	Length of the feed line	2
$L_f$	Length of the patch	3.3
$W_1$	Width of the patch	5.3
$L_1$	Length of the each slot	0.75
$W_1$	Width of the slot	0.5

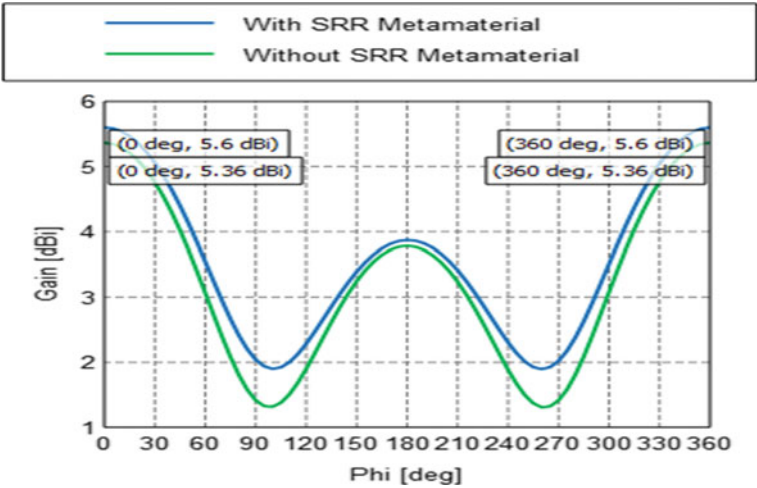


**Fig. 7** Graph of return loss and bandwidth in comparison with and without SRR based on metamaterial

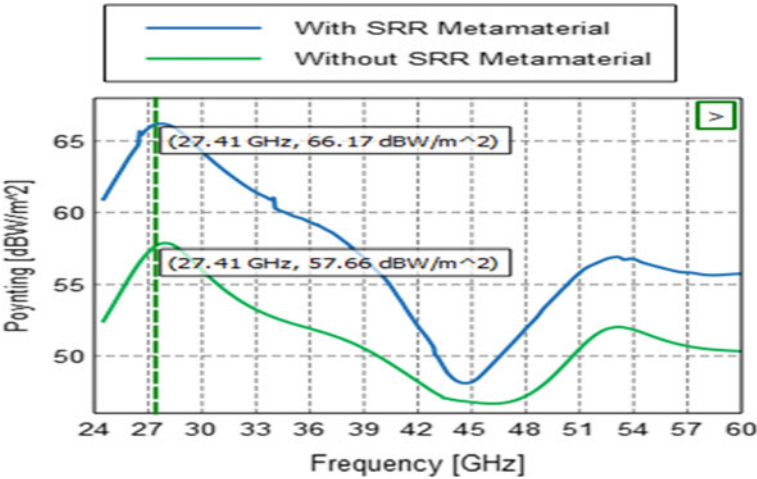
frequencies measured in Hz (Hertz). The typical requirements of return loss values for 5G applications are at least  $-10$  dB or more which satisfies the requirements of the proposed antenna, whereas the bandwidth is massive signifying very fast wireless communication of data which is better than [3, 6].

Figure 8 shows the graph of gain in dBi which signifies the directional capabilities and signal strength at 5.6 dBi with respect to  $\phi$  indicating strong power radiation for good 5G wireless communication.

Figure 9 shows the power density which signifies high-power radiation by the transmitting antenna per unit area in isotropic direction. It is also related to the



**Fig. 8** Graph of gain in comparison with and without SRR-based metamaterial



**Fig. 9** Graph of power density in comparison with and without SRR metamaterial

gain of the antenna. This power radiation density will help the receiving antenna to measure how much amount of radiated energy could be harvested per unit area.

Hence, from Figs. 7, 8, and 9, it is seen and investigated that the incorporation of split-ring resonator-based metamaterial enhances the performance of the proposed house-shaped antenna (Table 3).

**Table 3** Comparative performance analysis

Parameter	Without SRR	With SRR metamaterial	Frequency (GHz)
Return loss (dB)	−19.42	−21.79	27.5
Bandwidth (GHz)	4.770	5.371	27.5
Gain (dBi)	5.374	5.552	27.5
Power density (dBW/m <sup>2</sup> )	57.66	66.17	27.5

## 4 Conclusion and Future Research

The multi-band house-shaped microstrip patch antenna for 5G applications has been successfully developed, implemented, and investigated at 27.5 GHz using FEKO. It is therefore concluded that the metamaterial-based split-ring resonator (SRR) has peculiar electromagnetic properties such as simultaneous negative permeability and permittivity (negative refractive index) that are not commonly found in other natural materials, and this helps to boost the efficiency of the microstrip patch antenna in terms of improvement in return loss, bandwidth, gain, and power density. As a result, integrating the SRR-based metamaterial on the microstrip patch antenna, the bandwidth increased by 600 MHz, the return loss increased by  $-2.37$  dB, and the gain increased by 0.24 dBi, power radiation density increased by  $8.51$  dBW/m<sup>2</sup>. Future work can be done by designing a fractal patch antenna and shaping its array for enhancing the more return loss and gain and then integrating some complex formed metamaterial structure to increase the negative permeability and negative permittivity value through manipulation of flow of electromagnetic waves and further to improve its performance.

## References

1. Wojciech JK, Thanh NC (2019) Metamaterials in application to improve antenna parameters. In: Metamaterials and Metasurfaces. IntechOpen
2. Colaco J, Lohani R (2020) Design and implementation of microstrip circular patch antenna for 5G applications. In: 2nd international conference on electrical, communication and computer engineering. IEEE, Turkey
3. Mitra RN, Dharma PA (2002) 5G mobile technology: a survey. ICT Express Korean Instit Commun Inf Sci 2015:132–137
4. Basha AE, Huda AM, Zuhairiah ZA (2019) Reconfigurable radiation pattern of planar antenna using metamaterial for 5G applications. Materials 13(3):582
5. Lei T, Jianchun X, Li H, Yanan H, Shango H, Ming L, Ke B (2018) Bandwidth enhancement of microstrip patch antenna using complementary rhombus resonator. Wireless Commun Mob Comput 2018(6352181):8

6. Singh AK, Raman A (2018) Multiband microstrip patch antenna design for 5G using meta-material structure. In: 2nd international conference on trends in electronics and informatics (ICOEI). Tirunelveli, India, pp 909–914
7. Punith S (2020) A novel multiband microstrip patch antenna for 5G communications. *Procedia Comput Sci* 171:2080–2086
8. Weiren Z (2019) Electromagnetic metamaterial absorbers: from narrowband to broadband. IntechOpen
9. Colaco J, Lohani RB (2020) IoT based biomedical intelligent system in agriculture. In: International conference on iot based control networks and intelligent systems 2020. Elsevier, Kerala, India
10. Nader E, Richard WZ (2006) *Metamaterials: physics and engineering exploration* (Wiley) Book. IEEE
11. Pang Y (2020) Wideband RCS reduction metasurface with a transmission window. *IEEE Trans Antennas Propag* pp 7079–7087

# Family of Friends—A Hostel Utility System



**Mahima Badjate, Saniya Ponkshe, Sneha Rohida, Toral Patel,  
and Prajakta Deshpande**

**Abstract** The main objective and the driver force for building an android hostel management system is to increase the efficiency of how a typical hostel works while taking care of factors such as security, administration, and student portal. Parents who send their children to the hostels whether it is for education or for some different domestic reasons are concerned about their child's safety the most. Digitizing the whole process will make things easier to monitor. Being the students from information technology, we always have seen a lot of new applications being used in the college like the portal for result declaration, etc. The hostels still work in an old-fashioned manner where there goes a lot of manual energy in keeping up with the whole administration, that is where hostel management systems will be useful.

**Keywords** Android · Machine learning · ERP systems · Hostel management

## 1 Introduction

This is an android-based hostel management application. The need for automation in such type of an entity has been prolonged for really great time. There has been a lot of retrospection and thought over what a student desires when she enrolls into one and how can the student as well as the staff benefit from it. The needs and the specifications may differ from hostel to hostel, but we have tried to generalize

---

M. Badjate (✉) · S. Ponkshe · S. Rohida · T. Patel · P. Deshpande  
MKSSS's Cummins College of Engineering for Women, Karve Nagar, Pune 411052, India  
e-mail: [mahima.badjate@cumminscollege.in](mailto:mahima.badjate@cumminscollege.in)

S. Ponkshe  
e-mail: [saniya.ponkshe@cumminscollege.in](mailto:saniya.ponkshe@cumminscollege.in)

S. Rohida  
e-mail: [sneha.rohida@cumminscollege.in](mailto:sneha.rohida@cumminscollege.in)

T. Patel  
e-mail: [toral.patel@cumminscollege.in](mailto:toral.patel@cumminscollege.in)

P. Deshpande  
e-mail: [prajakta.deshpande@cumminscollege.in](mailto:prajakta.deshpande@cumminscollege.in)

**Table 1** Tabulated analysis of the literature survey

	Family of friends	Odoo OpenERP hostel management	C-ProGenial hostel management system	SabeeApp
Notice board	Present	Absent	Absent	Present
Parent SMS notification	Present	Absent	Absent	Absent
Complaint cell	Present	Absent	Present	Absent
Mess poll	Present	Absent	Absent	Absent
Two-step attendance	Present	Absent	Absent	Absent
Night out information	Present	Absent	Absent	Absent

the entire application for better scaling and usage. The application borrows many features from an already existing ERP systems for hostel management [1]. As of now, the application is free of cost unlike other established systems. As this is a very tedious job for the administrator, so automating this process will not only make the job easier but faster too. The use of an interactive user interface will not only attract the student users to it but make hostel life simpler for them. The idea is to treat every functionality as modules depicting the modularity. We have designed a logo and even named our application as Family of Friends—A Hostel Utility System.

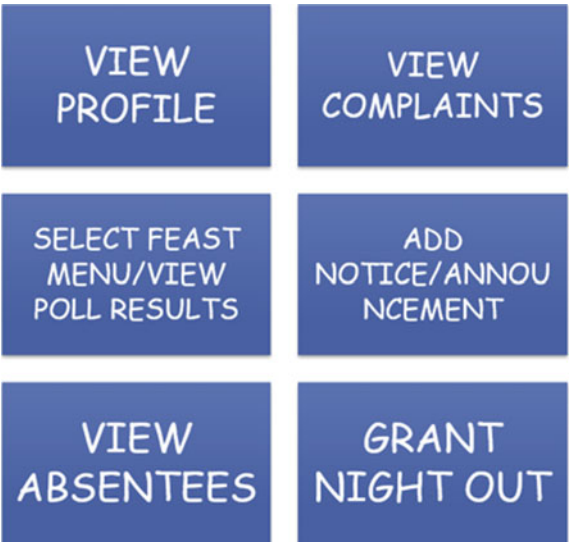
2 Literature Survey

The literature survey is simply a comparison between few exiting ERP systems which have hostel management as one of its sub-systems. The following is the detailed survey which gives a brief comparison based on very important parameters. Features like 2-step attendance, online night out form, SMS Alert to parent/guardian are something which makes FOF stand out and bridges the gap with the existing ERP systems. Most of the ERP systems are not exclusively for hostel management and do not completely focus on the same (Table 1).

3 Module Split-Up

The project is split into two major modules which are the student module and the admin module, wherein all the functionalities of both modules interact among them, and run the entire project as one module which runs as a wholesome entity (Fig. 1).

Fig. 1 Module split-up



3.1 Student Module

Module-I: Student registration:

The student will have to provide all sorts of details such as name, date of birth, address, blood group, medical history, allergies (if any), parent’s details, guardian’s details, etc., for registration process. Students can upload the documents as per their convenience. After submitting, a confirmation SMS will be sent to students *registered number and the student will be able to login with the username and password set* (Fig. 2).

Module-II: SMS alert for night out:

Fig. 2 Student module functionalities



If student wishes to apply for night out permission, then they will have to place a request for night out providing the necessary details such as the day and time he/she will return back to the hostel with the reason. As soon as they fill it, a notification will be sent to their parents number and this number would be taken during the registration itself. After confirmation from parents, the night out permission will be granted and parents will get to know their child's whereabouts. This adds up to the safety of the student.

### **Module-III: Notice board:**

Every student will have a dashboard like view on their respective accounts which will contain all sorts of important activities of the hostel. One such feature is the notice board where all the important announcements, notices, and circulars will be displayed. The admin has to just upload the information on the app, and the students can view at their end. If at all a notice has a deadline to be followed, it will also be mentioned, for example the payment of fees.

### **Module-IV: Attendance:**

One of the most talked about things about the attendance in hostels is that there is a whole lot of chaos while taking the attendance so the students have to mark the attendance from their room itself. It is a two-step verification where students have to first scan the QR code and each and every bed will have a separate QR code and after scanning the QR code, the IMEI number of the phone will be compared with the one taken during the registration, and once the IMEI number is matched, the attendance would be noted. Also, there is a specific time limit for attendance. Again, this adds up to additional safety.

### **Module-V: Mess poll:**

On a feast day at the hostel, there are polls conducted manually to let the students choose what to eat. This too can be given with the ease of their mobile phones. The admin will put up the potential menu that could be prepared by the kitchen staff prior to the feast day, and the students will have to poll for their dish, the one with maximum number of votes, wins!

### **Module-VI: Complaint and grievance cell:**

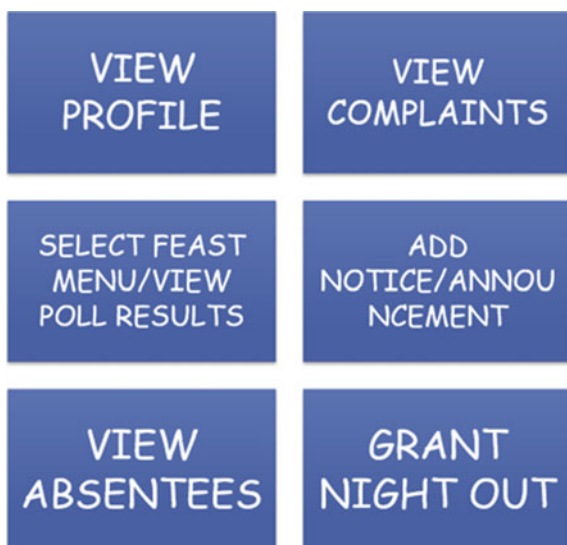
Complaints are to be heard and yet even today most of them are not yet acknowledged. There might be some students who are shy to complain, so here the students can complain by using their identity or anonymously. The students can complain at any time; as per their convenience also, complaints are prioritized so that quick action can be taken.

### **Module-VII: Feedback portal:**

For any organization to work properly, feedback is very important; hence, the dashboard will also give this feature for the students to give their valuable and honest feedback.



**Fig. 3** Admin module functionalities



### 3.2 Admin Module

#### Module-I: View profile:

All the student details filled during the registration process can be viewed by the admin. The admin can modify the details only if required (Fig. 3).

#### Module-II: View complaints:

It is the task of the admin to verify and acknowledge all the complaints issued by the students based on priority and convey it further to the appropriate faculty.

#### Module-III: Select feast menu:

Based on the conversation with the kitchen staff, admin puts together a list of food items which can be prepared on the feast day and the list is uploaded on to the dashboard, which the students poll for. The admin has to finalize the feast menu which has been selected by majority of students through the polls conducted and then based on maximum votes display the poll results.

#### Module-IV: Add notices:

The admin has to periodically update the notice board. The admin can add different notices and announcements based on the year or the major of the student.

#### MODULE-V: VIEW ABSENTEES:

The admin can keep track of the student's attendance by viewing absentees for that day. The admin has the record of students on night out. The students absent for 2 days will be notified on their mobile phones as well as their parent's mobile.

#### **Module-VI: Grant night out:**

Granting night out permission is the task of the admin. Once a student has requested for a night out permission, the admin looks into it, alerts the parent by sending an SMS, and then only grants permission for night out.

## **4 Tools/Technologies Used**

### **Android Studio**

Android Studio is standard and completely integrated development environment, which has been recently dispatched by Google for the android working framework. It has been intended to give new apparatus to application advancement.

### **MySQL Database**

As the data for the system is deliberately going to be structured, using MySQL database seemed a better option as it is best used in storing, manipulating, and retrieving data stored in a relational database.

### **MySQL Workbench**

It supports tables, views, triggers, etc., everything that is required in a database. MySQL workbench has a built in SQL visual editor which makes it very easy to manipulate the data. The visual SQL editor allows building, editing, and running queries against the MySQL server databases.

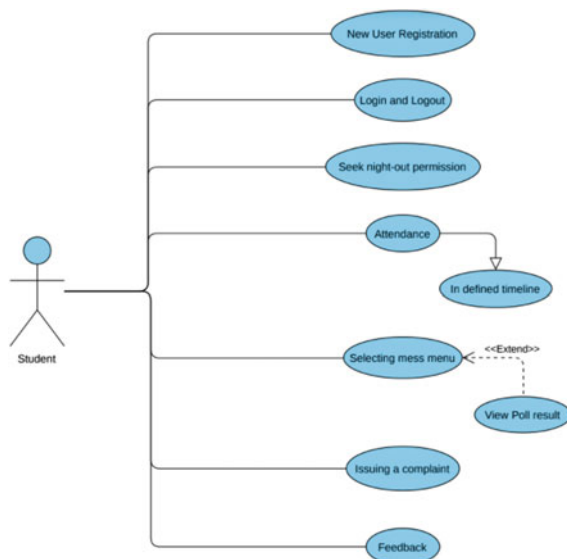
### **Bitrix24 tool for project management**

Bitrix24 is an assembled work space which handles the numerous parts of daily activities and assignments. Some of its project management highlight includes tasks, subtasks, conditions, Gantt, Kanban, recurring tasks, and task time-tracking.

### **GitHub for Project Collaboration**

GitHub Inc. is an auxiliary of Microsoft which gives facilitating to programming advancement and version control utilizing Git. It offers the dispersed version control and source code management functionality of GitHub, in addition to its own highlights.

**Fig. 4** Student-use case scenario



## 5 UML Diagram

In order for a software system to be of value, it should meet both functional requirements (FRs) and non-functional requirements (NFRs) [2]. UML diagram takes care of both the requirements. Given here is the student use case scenario (Fig. 4).

## 6 Deployment and Future Scope

The student module is very well ideated and implemented with all the functionalities running. The admin module is complete in terms of the required functionalities; yet, there is a scope to add a few features in the upcoming phase of the project. The app should be rendered via a server in order to use the application by the students. For this, we have chosen the college's local server to render the entire application which can be downloaded via a link on mobile phones.

Being free of cost, it can also scale up to a really large extent and can be made useful in various organizations which require such kind of a management system.

The data primarily being manageable, use of cloud is avoided but can be leveled up to cloud in future, if needed. The application is designed keeping in mind all the requirements by a manual run or not automated hostel management system.

This is one of a kind of a system where the entity chosen is the hostels; in the future, the system could be made generic for use for more such entities.

Room allocation feature can be added by making use of machine learning algorithms and then integrating with the system to even ease up this process.

## 7 Conclusion

The main motivation behind building this system is to automate the hostel management process. This system eases the load on the management team and simultaneously takes care about the security of the students. Key features which are added to support this are two-step attendance, SMS alert. They served to be one of the striking features to aid this motivation. Night out information to the parents via SMS alerts the parents about the whereabouts of their ward. This system also takes care about how easily the administrator is able to manage the whole process right from selecting the feast menu to taking the attendance right from their devices. Concluding, this is one of a kind of an application providing a wholesome user experience, easy to handle, and very simple logics which are well ideated to inculcate a lot of security aspect and is unique.

## References

1. Lu X, Nagelkerke M, Wiel DVD, Fahland D (2015) Discovering interacting artifacts from ERP systems. In: IEEE transactions on services computing, vol. 8, no. 6, pp. 861–873. <https://doi.org/10.1109/TSC.2015.2474358>
2. Supakkul LCS (2005) A UML profile for goal-oriented and use case driven representation of NFR's and FR's. In: Third ACIS int'l conference on software engineering research, management and applications (SERA'05)

# Vulnerability Due to Training Order in Split Learning



Harshit Madaan, Manish Gawali Viraj Kulkarni, and Aniruddha Pant

**Abstract** Split learning (SL) is a privacy-preserving distributed deep learning method used to train a collaborative model without the need for sharing of patient's raw data between clients. In split learning, an additional privacy-preserving algorithm called no-peek algorithm can be incorporated, which is robust to adversarial attacks. The privacy benefits offered by split learning make it suitable for practice in the healthcare domain. However, the split learning algorithm is flawed as the collaborative model is trained sequentially; i.e., one client trains after the other. We point out that the model trained using the split learning algorithm gets biased toward the data of the clients used for training toward the end of a round. This makes SL algorithms highly susceptible to the order in which clients are considered for training. We demonstrate that the model trained using the data of all clients does not perform well on the client's data which was considered earliest in a round for training the model. Moreover, we show that this effect becomes more prominent with the increase in the number of clients. We also demonstrate that the SplitFedv3 algorithm mitigates this problem while still leveraging the privacy benefits provided by split learning.

**Keywords** Split learning · Distributed deep learning · Federated learning

## 1 Introduction

In the healthcare sector, artificial intelligence can help predict a patient's medical condition using deep learning models. A deep learning model becomes more accurate and robust when it is trained on a large amount of diverse data [5]. It would be ideal if multiple hospitals can cooperate and share their patient's raw data to train a single model. However, laws concerning data privacy like GDPR, HIPAA, and DISHA impede hospitals from sharing patient's raw data with any authority. Distributed

---

H. Madaan (✉)

Indian Institute of Science Education and Research, Pune, India

e-mail: [I.harshit.madaan@students.iiserpune.ac.in](mailto:I.harshit.madaan@students.iiserpune.ac.in)

H. Madaan · M. G. V. Kulkarni · A. Pant

Deeptek Inc., Pune, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

103

M. Tuba et al. (eds.), *ICT Systems and Sustainability*, Lecture Notes in Networks and Systems 321, [https://doi.org/10.1007/978-981-16-5987-4\\_11](https://doi.org/10.1007/978-981-16-5987-4_11)

deep learning methods enable a model to be trained by the collaboration of multiple clients and without compromising the privacy of these clients. For example, In the federated learning algorithm, a client only shares the model parameters to a central server instead of the patient's raw data. Three privacy-preserving distributed deep learning methods are described in the following subsections.

### ***1.1 Federated Learning***

In federated learning [9–11], the clients train in parallel for a specified number of rounds (or global epochs). During a round, the global model is sent by the server to the clients. Each client that gets the copy of a global model trains it for some local epochs with their respective datasets; thus, each copy of the global model is updated in a different way by different clients. These models are sent back to the server and are then aggregated and averaged on the server. The averaged model is sent back to the clients for validation. This process concludes one round (global epoch) of training.

### ***1.2 Split Learning***

In split learning [4, 13], the model architecture is divided into two parts, the front part of the model resides on the client and the rest of it is on the server. The model is trained sequentially among clients in split learning algorithms. During a round, forward propagation is carried out on the part of the model that resides on the client end, and the last layer's activations are sent to the server where the rest of the model resides. The activations received are used for forward propagation on the server-side model, after which backpropagation is done on the server-side model, and the necessary gradients are sent to the client where backpropagation is carried out on the client-side model. After that, the next client trains the model by communicating with the server, following the same trend of sending activations, receiving gradients, and updating the model. One round (or global epoch) of training completes when all the clients have trained by communicating with the server. Although the server side of the model is the same for every client, there is no notion of the global model in SL as the part of the model architecture that resides on the client side is unique for each client. If there are 'n' clients, then there are 'n' different client-side models and a single server-side model.

### ***1.3 SplitFed***

SplitFed is a hybrid approach between split learning and federated learning. There are two variants of SplitFed proposed by Thapa et al. [16], namely SplitFedv1 and

SplitFedv2, and a recent SplitFed approach termed as SplitFedv3 proposed by Gawali et al. [3]. In SplitFed algorithms, the model architecture is divided into segments similar to split learning. The training happens parallelly in SplitFedv1 and SplitFedv3, whereas a model is trained sequentially in SplitFedv2. In SplitFedv2, averaging of the client-side model takes place at the end of each global epoch. In SplitFedv3, the server-side model is averaged out at the end of a global epoch. In SplitFedv1, both parts of the model, i.e., client-side and server-side, are averaged out at the end of a global epoch.

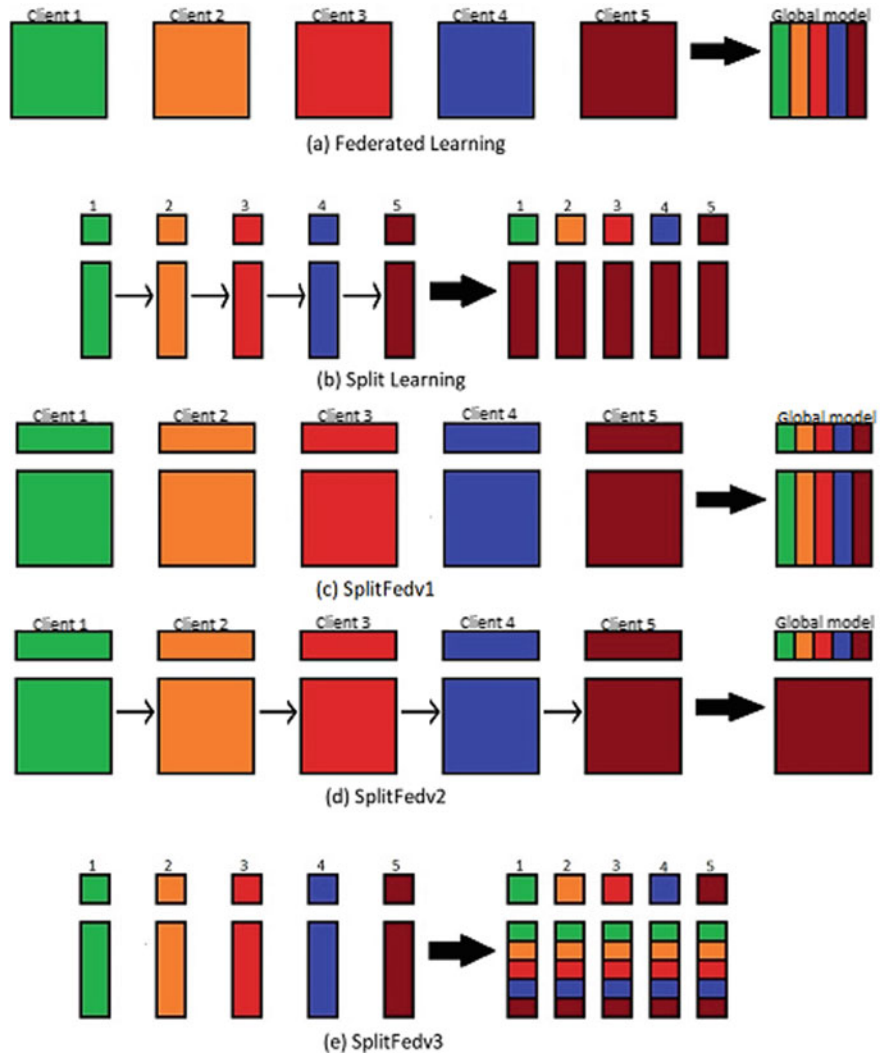
## 2 Related Work and Vulnerability in SL

When clients collaborate to train a deep learning model, the model can be trained either sequentially (i.e., one client trains after the other) or parallelly (all the clients train at the same time). When a collaborative model is trained sequentially, it can lead to forgetting when the training set comprises data from non-iid data sources. Each time the model is updated by the data of a client, the model parameters will be updated in such a way to favor the recently seen data. The updated model might forget what was learned when it had trained with the previous client's data resulting in poor classification performance on the 'same' dataset. This was shown by Sheller et al. [15] for distributed deep learning methods like Institutional Incremental Learning (IIL) and Cyclic Institutional Incremental Learning (CIIL), where the model is trained sequentially (Fig. 1).

In split learning, even though the client-side model is unique for each client, a large trainable portion of the model resides at the server end, and the server-side model is trained sequentially. Forgetting can manifest in split learning such that the performance of the model on the clients that are training early in a cycle could be inadequate.

In the case of federated learning, the global model is an average of all the local models with each client. Each client would have an equal contribution to the final state of the global model when the number of data samples for all clients are the same; hence, there is no possibility of forgetting in FL. And even when the number of data samples is uneven, there would still be some contribution from each client.

For the Split-fed approach, SplitFedv2 might also suffer from the problem of forgetting as it also follows the pattern of cyclical training. SplitFedv1 and SplitFedv3, on the other hand, will not show forgetting as the server side is averaged and does not follow a cyclic pattern. SplitFedv1 has an additional communication overhead where we send all the client-side models to a place where the averaging happens.



**Fig. 1** Picture the boxes being the state of the model when the model has trained on the respective client dataset. Due to the cyclic pattern of training in split learning, the server-side model's final state is heavily molded according to the requirements of the clients that were considered for training toward the end of the training loop and can result in forgetting. Whereas in a federated setting, the final model is an average of the local models, and thus, every client collaborates equally



### 3 Data and Experimental Settings

This section describes the datasets and experimental settings for distributed deep learning methods.

#### 3.1 Data

We aggregated the data from three private hospitals (referred to as A, B, C) and two publicly available datasets (MIMIC [7] and Padchest [1] referred to as D, E, respectively). The data contains chest X-rays with healthy and TB-suspected images of resolution  $224 \times 224$  pixels. Each chest X-ray is labeled either TB-suspect or TB-negative by a team of expert radiologists. The training set had a total of 8708 X-ray images. In addition to the training set, the validation and testing set contained 2500 images each. The prevalence (percentage of TB-suspect images) of the training set is 50 and 10% for the validation set and test set (Table 1).

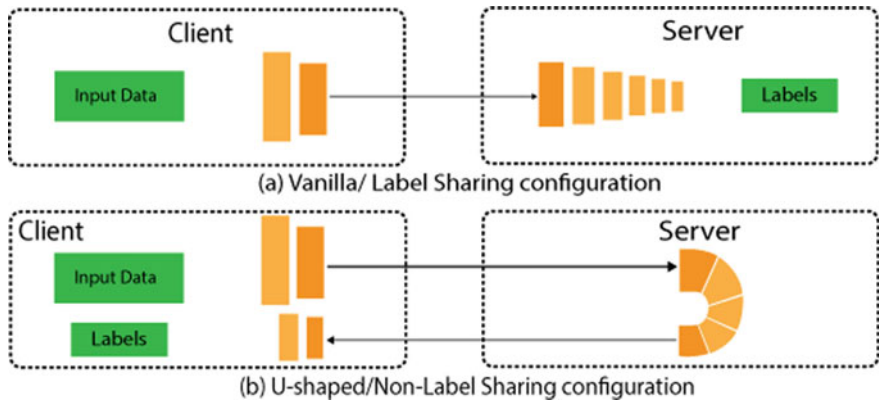
#### 3.2 Experimental Settings

We used Pysyft [14] library for the experimental implementations. The hospital settings were created using virtual workers, and each virtual worker has access to its own data with no communication to other workers. The Densenet-121 [6] architecture was used for classification with an Adam optimizer [8] and a learning rate of  $1e-4$ . A combination of threshold dependent (F1-score and Cohen’s kappa [2]) and threshold independent (Auprc) metrics are used to evaluate the saved models. We did not consider auc-roc metric due to the prevalence of 10% in the test set [12]. The threshold-dependent metrics are evaluated on a sensitivity of 0.81. The model with the least validation loss is saved for evaluation. We observed that the model converged within five epochs, and thus, each simulation was executed for ten epochs.

There are many SL configurations [17], which can also be extended to SplitFed. The most basic of all is the vanilla configuration. In the vanilla configuration, the clients have to share the labels of the data with the server. A more private configuration called a U-shaped configuration allows the data and the labels to reside on the client. We have used the U-shaped configuration for SL and SplitFed experiments (Fig. 2).

**Table 1** Distribution of chest X-ray images

Data	A	B	C	D	E	Total
Train	1816	3772	1150	880	1090	8708
Validation	500	500	500	500	500	2500
Test	500	500	500	500	500	2500



**Fig. 2** In a U-shaped configuration, the input data and labels reside on the client, enhancing privacy over a vanilla configuration where labels have to be shared with the server

**Table 2** Split learning results for every client

Client	AUPRC			F1-score			Kappa		
	First	Last	% drop	First	Last	% drop	First	Last	% drop
A	0.4318	0.5833	25.97	0.26	0.41	36.59	0.1107	0.3032	63.49
B	0.4915	0.8651	43.19	0.30	0.76	60.53	0.1607	0.7340	78.11
C	0.5738	0.7940	27.73	0.53	0.74	28.38	0.4528	0.7064	35.90
D	0.4900	0.6897	28.95	0.55	0.63	12.70	0.4794	0.5788	17.17
E	0.7258	0.8244	11.96	0.71	0.78	08.97	0.6787	0.7510	09.63

The second column is the auprc score when that respective client is placed first in training, and the third column is the auprc score when that client is placed at the end of the training loop. The fourth column notes down the percentage change of the performance of the client training early compared to training at the end of a global epoch. A similar pattern is followed for Kappa and F1-scores

4 Experiments and Results

The first set of experiments compared the performance changes of a client for two extreme permutations of client order training among the different distributed deep learning algorithms, namely federated learning, split learning, and SplitFedv3. We have considered two cases for each client, wherein for the first case, a client was the first client, and for the second case, that same client trained at the end during a global epoch. For every run, the F1-score, Cohen’s kappa, and auprc scores were calculated. The training loop has all the clients; just the permutation of the clients is different for different runs. Tables 2, 3, and 4 show the results for split learning, federated learning, and SplitFedv3, respectively.

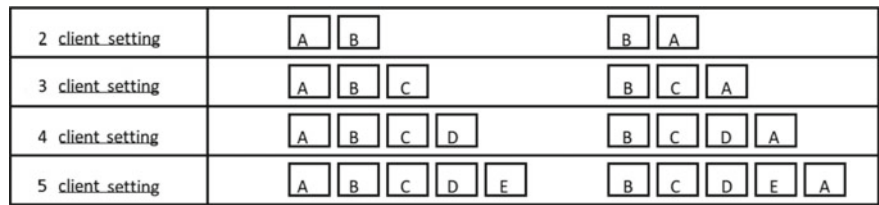
As seen from Table 2, in split learning, there was a significant performance drop in all the three performance metrics for each client depending on its order of training. On average, a percentage change of 31% was observed across all the metrics, with

**Table 3** Federated learning results for every client

Client	AUPRC			F1-score			Kappa		
	First	Last	% drop	First	Last	% drop	First	Last	% drop
A	0.4482	0.4482	–	0.35	0.35	–	0.2327	0.2327	–
B	0.7646	0.7646	–	0.60	0.60	–	0.5392	0.5392	–
C	0.7844	0.7844	–	0.68	0.68	–	0.6428	0.6428	–
D	0.7189	0.7189	–	0.66	0.66	–	0.6096	0.6096	–
E	0.7656	0.7656	–	0.71	0.71	–	0.6787	0.6787	–

**Table 4** Splitfedv3 results for every client

Client	AUPRC			F1-score			Kappa		
	First	Last	% drop	First	Last	% drop	First	Last	% drop
A	0.4543	0.4543	–	0.34	0.34	–	0.2171	0.2171	–
B	0.8063	0.8963	–	0.63	0.63	–	0.5788	0.5788	–
C	0.6071	0.6071	–	0.60	0.60	–	0.5454	0.5454	–
D	0.6731	0.6731	–	0.67	0.67	–	0.6226	0.6226	–
E	0.8060	0.8060	–	0.75	0.75	–	0.7175	0.7175	–



**Fig. 3** Client settings that were considered for the experiments

the maximum percentage change being 78% in the kappa score for client A. Drastic changes were observed for client A as all the three metrics record a percentage change higher than 40%. Tables 3 and 4 show that client order does not matter in federated learning and splitfedv3 as there was no performance drop observed between training early and training late for any client. The clients train the model parallelly in splitfedv3 and federated learning, which makes them immune to any ordering issues (Fig. 3).

The next set of experiments tracks the performance of a single client (client A) over different client settings. For each client setting, the simulations were run for two extreme permutations (one where client A was in the beginning and the other where it trained at the end), and the metric scores of client A were recorded. Initially, only two clients were considered in the training loop, and then the clients were added incrementally one at a time to the training loop. A total of four different client settings were designed for the comparison of performance metrics between the various distributed deep learning algorithms (Table 7).

**Table 5** Split learning results for client A over different settings

Client A	AUPRC			F1-score			Kappa		
	First	Last	%drop	First	Last	%drop	First	Last	%drop
2 client setting	0.5986	0.6616	09.52	0.44	0.51	13.72	0.3414	0.4281	20.25
3 client setting	0.5099	0.6735	24.29	0.37	0.48	22.92	0.2604	0.3905	33.32
4 client setting	0.3750	0.5928	36.74	0.31	0.39	20.51	0.1806	0.2816	35.87
5 client setting	0.4318	0.5833	25.97	0.26	0.41	36.59	0.1107	0.3032	63.49

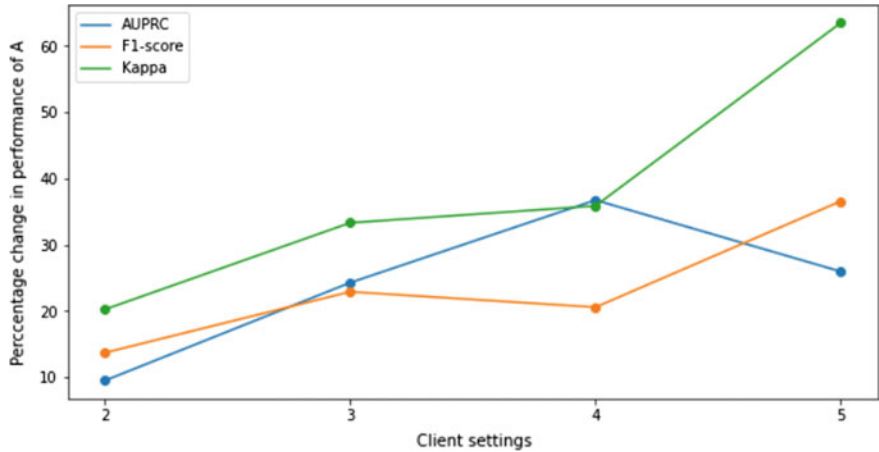
**Table 6** Federated learning results for client A over different settings

Client A	AUPRC			F1-score			Kappa		
	First	Last	% drop	First	Last	% drop	First	Last	% drop
2 client setting	0.6198	0.6198	–	0.44	0.44	–	0.3506	0.3506	–
3 client setting	0.5041	0.5041	–	0.42	0.42	–	0.3154	0.3154	–
4 client setting	0.3012	0.3012	–	0.27	0.27	–	0.1272	0.1272	–
5 client setting	0.4482	0.4482	–	0.35	0.35	–	0.2327	0.2327	–

**Table 7** SplitFedv3 Learning results for client A over different settings

Client A	AUPRC			F1-score			Kappa		
	First	Last	% drop	First	Last	% drop	First	Last	% drop
2 client setting	0.6132	0.6132	–	0.45	0.45	–	0.3601	0.3601	–
3 client setting	0.5118	0.5118	–	0.43	0.43	–	0.3354	0.3354	–
4 client setting	0.4734	0.4734	–	0.38	0.38	–	0.2650	0.2650	–
5 client setting	0.4543	0.4543	–	0.34	0.34	–	0.2171	0.2171	–

Tables 4, 5, and 6 show the result of the effect of client settings in the distributed deep learning algorithms. As shown in Fig. 4, there was an increase in the performance difference of the first client in split learning as more and more clients were added. This shows that forgetting becomes more prominent with an increase in the number of clients. No such performance difference was observed in Federated learning and SplitFedv3. A general decrease in performance was observed for the first client across all the deep learning methods. Although both SplitFedv3 and Federated learning were stable over the ordering changes, the performance results of splitFedv3 were better compared to federated learning for any client setting across all three metrics.



**Fig. 4** Graph of percentage change vs. client settings for split learning results. The x-axis is the number of clients in the particular setting. The y-axis is the percentage change in the performance of client A when its order in the cycle changes from being in the beginning to being at the end

5 Conclusion

Split learning is prone to client-ordering, which is a repercussion of training the collaborative model sequentially. Our results showed that for the split learning algorithm, the performance of the trained model on the data of the client that is placed at the beginning of the training order is compromised while testing. We also showed that the drop in the performance of a client encountered in split learning is absent in other distributed deep learning methods like federated learning and SplitFedv3 due to parallel training between the clients that are collaborating to train a single model. Further, we showed that in split learning, the difference between the test performance of a client when training at the earliest and when training at the end of a training loop worsened with an increase in the number of clients in between. SplitFedv3 has a parallel training pattern making it immune to client-ordering issues. It also requires low computation resources on the client side as only a part of the model is trained on the client side. Moreover, split learning algorithms for privacy like no-peek can also be adopted in SplitFedv3.

References

1. Bustos A, Pertusa A, Salinas JM, de la Iglesia-Vayá M (2020) Padchest: a large chest x-ray image dataset with multi-label annotated reports. *Med Image Anal* 66(101):797
2. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20(1):37–46

3. Gawali M, Suryavanshi S, Madaan H, Gaikwad A, KN BP, Kulkarni V, Pant A, et al (2020) Comparison of privacy-preserving distributed deep learning methods in healthcare. arXiv preprint [arXiv:2012.12591](https://arxiv.org/abs/2012.12591)
4. Gupta O, Raskar R (2018) Distributed learning of deep neural network over multiple agents. *J Netw Comput Appl* 116:1–8
5. Hestness J, Narang S, Ardalani N, Diamos G, Jun H, Kianinejad H, Patwary M, Yang Y, Zhou Y (2017) Deep learning scaling is predictable, empirically. arxiv 2017. arXiv preprint [arXiv:1712.00409](https://arxiv.org/abs/1712.00409) (2017)
6. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 4700–4708
7. Johnson AE, Pollard TJ, Greenbaum NR, Lungren MP, Deng CY, Peng Y, Lu Z, Mark RG, Berkowitz SJ, Horng S (2019) Mimic-cxr-jpg, a large publicly available database of labeled chest radiographs. arXiv preprint [arXiv:1901.07042](https://arxiv.org/abs/1901.07042)
8. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)
9. Konečný J, McMahan HB, Ramage D, Richtárik P (2016) Federated optimization: distributed machine learning for on-device intelligence. arXiv preprint [arXiv:1610.02527](https://arxiv.org/abs/1610.02527)
10. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA (2017) Communication-efficient learning of deep networks from decentralized data. In: *Artificial intelligence and statistics*. PMLR pp 1273–1282
11. McMahan B, Rampage D Federated learning: collaborative machine learning without centralized training data. <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>
12. Ozenne B, Subtil F, Maucourt-Boulch D (2015) The precision-recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases. *J Clin Epidemiol* 68(8):855–859
13. Poirot MG, Vepakomma P, Chang K, Kalpathy-Cramer J, Gupta R, Raskar R (2019) Split learning for collaborative deep learning in healthcare. arXiv preprint [arXiv:1912.12115](https://arxiv.org/abs/1912.12115)
14. Ryffel T, Trask A, Dahl M, Wagner B, Mancuso J, Rueckert D, Passerat-Palmbach J (2018) A generic framework for privacy preserving deep learning. arXiv preprint [arXiv:1811.04017](https://arxiv.org/abs/1811.04017)
15. Sheller MJ, Reina GA, Edwards B, Martin J, Bakas S (2018) Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In: *International MICCAI brainlesion workshop*. Springer, pp 92–104
16. Thapa C, Chamikara MAP, Camtepe S (2020) Splitfed: when federated learning meets split learning. arXiv preprint [arXiv:2004.12088](https://arxiv.org/abs/2004.12088)
17. Vepakomma P, Gupta O, Swedish T, Raskar R (2018) Split learning for health: Distributed deep learning without sharing raw patient data. arXiv preprint [arXiv:1812.00564](https://arxiv.org/abs/1812.00564)

# Efficient Attendance Management System Based on Face Recognition



Ujwalla Gawande, Pratyush Joshi, Sumedh Ghatwai, Shreyas Nemade, Soham Balkothe, and Nishant Shrikhande

**Abstract** Every education institution nowadays is concerned about student attendance and performance. In the current academic system, consistent class attendance of students plays a significant role in their performance, assessment, and quality monitoring. The traditional approach of taking attendance in several institutions is by calling the names of students one after the other or each student manually signs on the papers. In existing approaches, taking and tracking student's attendance manually, losing attendance sheets, dishonesty of students, and high error scales are open challenges face by the faculties. It is a complex process, requires time, and causes manual paperwork. We proposed an efficient and robust approach for an attendance monitoring system using the face of a human. The proposed algorithm first detects and recognizes the student's faces from videos or images. Second, mark the attendance using a neural network model and texture features. Experiments were conducted on the commercially available datasets. The proposed approach is compared with the traditional attendance marking system in terms of time and accuracy. Our approach resolves the state-of-the-art approach challenges and saves time to monitor the students.

**Keywords** Face recognition · Automated student attendance system · OpenCV · Neural network

## 1 Introduction

In this modern era of automation, many scientific advancements and inventions have taken place to increase accuracy and to improve our lives. In a class, maintaining the attendance of students is important for checking their performance. The consistency of student attendance represents the discipline and is not likely to engage in destructive activities in classroom and other extracurricular and co-curricular academic activities [1, 2]. More absenteeism increases the risk of student failure

---

U. Gawande (✉) · P. Joshi · S. Ghatwai · S. Nemade · S. Balkothe · N. Shrikhande  
Department of Information Technology, Yeshwantrao Chavan College of Engineering, Nagpur,  
Maharashtra 441110, India

and unexpected abandon from academics [3, 4]. Every institute has its method in this regard. Currently, many faculties are taking attendance on paper and manually feeding attendance in the attendance of cell software. Some organization has adopted an automatic attendance mechanism using face, iris, and palm-based biometric techniques. These systems are widely used in different organizations. All the same, it is a time-consuming process of recording attendance and marking it to database, whereas in biometric systems. The conventional approach to monitor student attendance is not reliable and secure. It causes many issues such as it utilized a lot of time of lecture hours. Again, the proxies' attendance rate is high. In recent times, several student attendance monitoring systems have been introduced for resolving the complication of student attendance. The various biometric templates can be used for identification and validation. Automation of attendance system has an edge over traditional methods as it saves time and also can be used for security purposes. In [5] proposed, a Windows-based system in which students manually feed the attendance. If the student is present, then enroll data into the system. The limitations of this approach are (1) Again, attendance required time to take and manual. Instead of faculty students are marking their attendance. (2) Increase chances of false marking of attendance and security. In [6] proposed, a Bluetooth-based attendance system. In this approach, handheld mobile was used for the attendance using Bluetooth. The automatic push notification is generated on the student's phone. The limitations of this approach are (1) every student must have android mobile phones. (2) Bluetooth connectivity and synchronization, distance constraints affect the performance. In [7] proposed, a model for organization employees which is based on the fingerprint of employees used for the attendance system. This system compares the user fingerprint and the database enrolled user fingerprint sample for the attendance authentication. Limitations of this approach include (1) every classroom requires a fingerprint sensor for taking attendance, time requires recording the biometric captured information in the database. However, it is a secure approached but time-consuming and costly. In [8] describe the unique approach of taking attendance using palm recognition. However, these systems have limitation such as ease of use, combability, robustness, and authenticity. So, an attempt to conquer the shortcomings of the existing systems leads to the usage of face biometric for the attendance monitoring system. The face recognition is more authentic, secure, and robust as compared to other physiological and behavioral biometric trails such as iris, palm, and signature. Each human is having a unique facial feature. It is difficult to be faked or forged. Furthermore, the class or school teacher's uncertain student count issue gets resolved with the proposed approach. The proposed work's main contributions include (1) neural network-based approach to face detection and recognition from videos. (2) Surveillance camera's videos based on embedded systems for attendance in classrooms. These systems can be used for identifying the student and marking attendance. In this work, we have used platforms like OpenCV that makes the system more robust and secure, and useful for the real-time implementation of the attendance management system. The paper is organized as follows. Section 1 describes the attendance management system and its application. Section 2 describes the existing attendance monitoring system available in the literature. Section 3 represents the proposed framework. Section 4



represents the experimental results along with the graphical user interface of the proposed system. Finally, the paper concludes with the future research direction in Section 5 regarding the attendance management system.

## 2 Literature Survey

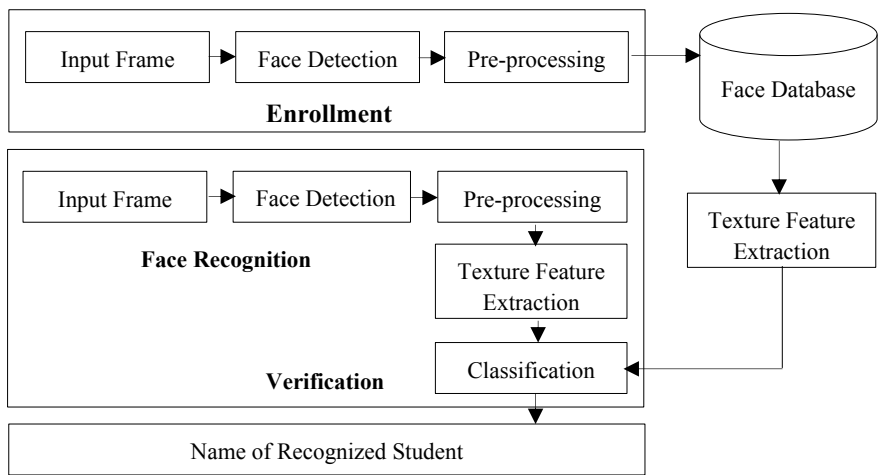
Traditionally token-based approaches such as key, magnetic card, PIN, and password are used for identity verification and have achieved more reliable verification and identification well known as biometrics [9, 10]. One of the most used and natural biometric is face recognition. For most of the existing developed surveillance cameras in public places, researchers tend to developed biometrics applications to achieve a higher degree of security. The biometric-based attendance system must be unique, secure, portable, and highly accurate. In recent time, face recognition is used by most of the organization due to the convenience and high accuracy compared to other biometric and non-biometric traits. The biometric-based approaches for the attendance management system are mostly used due to its advantages such as security, portability, and less costly. In [11] proposed, a FaceNet. In this approach, the information is divided into different clusters using the clustering technique. The different distance measurement metrics are used for such as the Euclidean distance matrix is used for face recognition. The similarity measurement matrix was useful in the classification of the face in the output. In [12] proposed, the use of principal component analysis (PCA) with ANN for face recognition in attendance management. In [13] approach utilized a fingerprint trait for attendance. A portable fingerprint scanner is used for enrolling the student information in the database. Each student enrolls fingerprint sample for marking their attendance in the system without the interruption of the instructor. The system is secure and handy for students as well as an institution. The issue with this approach is transferring the device from one student to another and the time required for taking the attendance. Sometimes the fingerprint sensor distracts the student's attention in the classroom. In [14] proposed, an attendance system framework is based on RFID. In this model, each student maintains their RFID-based identity cards. The problem with this approach is that any student can take any person's RFID card and registered the attendance. Unauthorized access to the card cannot be prevented that results in an absent student can be marked as present in this approach. The other traits such as the iris-based attendance framework were proposed by [15]. In [16] describe a secure and authenticated real-time face recognition-based system. It can be used in real time for face recognition. PCA is developed to overcome expensive computation and the need for great amounts of storage of older face recognition techniques such as correlation methods. PCA approach can be used to reduce the feature vector size. In the proposed framework, we have utilized this approach for face recognition of students in real-time environments. Due to existing system issues and challenges in real-time environment, it is observed that there is need of efficient and secure attendance management system to handle the existing attendance management limitations.

### 3 Proposed Methodology

The proposed system architecture is divided into two phases as (1) enrollment phase and (2) verification phase. In the first phase, each student register in the system with their basic information such as name, section, semester, year, branch, college, and face sample images. For each student, five images are recorded at the time of registration. In this process, each students’ distinct features are extracted using a texture feature extraction algorithm, and later at the time of the verification phase, these features are again extracted in real time to match with the features that are recorded at the time of registration as shown in Fig. 1. Both the phases are described in brief as follows.

#### 3.1 Input Capture

The camera is mounted at a distance from the entrance of the classroom to capture the front-facing images of the students. The surveillance captures the video’s later it gets converted into frames for further processing. Each extracted frame from a video is of the size  $1040 \times 768$  to avoid resizing of the image in the preprocessing.



**Fig. 1** Proposed automated attendance management system architecture

### **3.2 Face Detection**

In the literature, there are several face detection algorithms such as face geometry-based techniques, feature invariant-based methods, machine learning-based methods [17, 18]. In [16], framework gives a high detection rate and is also fast while detecting the faces from images or frames. It can be used in real time because it is fast and robust. In this algorithm, the integral image and AdaBoost classifier learning algorithms are used for face detection. This algorithm is efficient and robust in different lighting conditions. Again, we have used multi-resolution Haar classifiers to achieve a higher face detection rate [19].

### **3.3 Preprocessing**

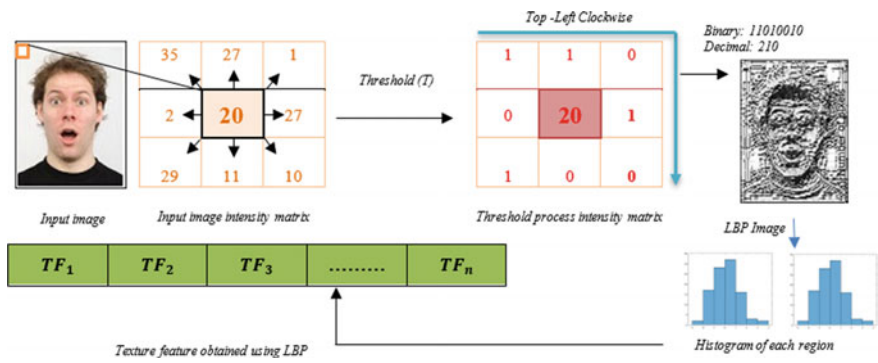
In preprocessing first, we have used histogram equalization of the extracted face images for contrast enhancement as it normalized the range of intensity values of the images. The histogram equalized image is clearer, and the intensity of the images is uniformly distributed over the range of intensity values of the input face images.

### **3.4 Database Development**

The student's face images are captured in the enrollment stage. Then each face image's features are extracted from the enhanced face images. Later these features can be used in the classification. The features are store in the database for further processing in the classification stage as shown in Fig. 1. In the enrollment stage, each face was captured in a different orientation and with different expressions. The various illumination conditions are maintained while capturing the face images of individuals.

### **3.5 Texture Feature Extraction Using Local Binary Pattern**

Each student's face is having a different unique texture feature. However, texture features are extracted for identifying each student. The texture information was extracted using local binary pattern (LBP) [17]. LBP mask operates on the input image, and after rotation over the complete image generates the decimal value for the center element as shown in Fig. 2. A pixel in the  $3 \times 3$  mask is compared with the neighborhood element in a clockwise direction that results in binary bits 0 and 1. For each block, binary pattern is concatenated from the top-left corner. The given pixel is replaced with the corresponding decimal values obtain after concatenation



**Fig. 2** Example of local binary pattern (LBP)

of binary bits. The histogram of each block is computed and concatenated to obtain the full image histogram.

In the histogram, each decimal bit is plotted. Each face image LBP pattern and the normalized histogram are different for individuals. The texture feature vector size after histogram normalization is  $1 \times 256$ . The face dataset consists of 100 student different face images. Each individual is having five different images, which are obtained in the enrollment phase. So total texture features are  $500 \times 256$ . After texture feature, extraction processes are further using a support vector machine classifier.

### 3.6 Post-Processing

In the post-processing, after recognizing the faces of the students, the names are updated into a text file. The SVM classifier output is the index label of students whose faces are matched from the database. The labeling of classes is mapped in the database for the specific student face features, respectively. Next, we have described the face recognition process in brief as follows

**Algorithm 1: Face recognition algorithm for attendance Management System**


---

```

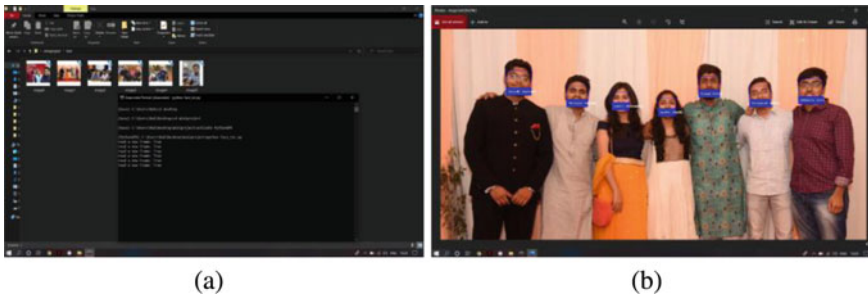
Step 1   Capture the video of student in a classroom
Step 2   Extracting the frames from the video
Step 3   Apply Viola-Jones algorithm for face detection
Step 4   Extract the region of interest and bounding box parameter
Step 5   Convert it into gray scale image
Step 6   Apply histogram equalization
Step 7   Process each individual enhance frame of video
         for  $f_i \leftarrow 1$  to  $n$  do
             extractfeat  $\leftarrow$  Texture_Feature_Extractor_LBP( $f_i$ )
             trainNN  $\leftarrow$  SVMtrain(enrollfeat)
             classLbl  $\leftarrow$  SVMTest(trainNN, extractfeat)
             if classLbl == matchResult then
                 student  $\leftarrow$  present
                 studName  $\leftarrow$  refList[classLbl]
             else
                 student  $\rightarrow$  absent
                 studName  $\leftarrow$  refList[classLbl]
             end
         end
Step 8   Display and write present and absent student details.
Step 9   End

```

---

## 4 Experimental Results

All the experiments were conducted on the publicly available dataset. The practical implementation of the proposed system is the OpenCV. Figure 3 shows the results of the proposed system where the multiple faces are detected accurately. Extracted frames from a video are represented in Fig. 3a. The qualitative results are shown in Fig. 3b. After classification using SVM, the names of recognized faces are written in the text document as shown in Fig. 3c. Hence, the proposed system accurately detects the faces and recognized the students and manages the attendance efficiently.



**Fig. 3** a Extraction of frames, b sample frame of recognized faces, c total present and absent names in all the extracted frames

**Table 1** Performance comparison of proposed approach and existing system

Type of feature extraction algorithm	Classification approach	Error rates		Detection rate (%)	TAR (%)	Training Time	Testing Time (per image)
		FAR (%)	FRR (%)				
Proposed Texture Feature + Viola jones [16]	SVM	1.7	1.9	96.8	97	0.903	0.241
LBP codes [7]	SVM	1.4	2.1	95	92	0.903	0.67
Color feature extraction [18]	Haar Cascade	1.7	1.2	92	94	0.903	0.56

The performance measurement matrix used for the evaluation of proposed system accuracy was False Match Rate (FMR), False Non-match rate (FNMR), and True Match Rate (TMR). Total 300 images are used for training, and 200 images are used for testing out of 500. The most promising result of 1.7% FAR and 1.9% FRR and recognition rate of 96.8% achieved by the SVM classifier. TAR faired to 97% and required a very less training time of 0.903 s. and a testing time of 0.241 s. per image as shown in Table 1. The proposed system accuracy is comparatively better than the existing system [7, 7], and the testing time required is less as compared to the state-of-the-art approaches. Table 1 shows that the proposed approach computationally less complex than the [7, 18]. In these filtering and preprocessing increase the time to before face detection. In [7] approach, contrast stretching, bilateral filtering, linear blending increase time to processed the input image. The proposed approach, histogram equalization and normalization, uniforms the intensity of the input image and does not require extra preprocessing and filtering steps before face detection and for texture feature extraction. In [18], approach uses the Haar cascade classifier. It decomposes the input image into multiple resolutions. Later, high- and low-frequency components are separated for further processing. The input image is process block-by-block at a different resolution. Hence, these extra processing steps increase the time and computational complexity required for the detection of legitimate students. The proposed approach is simple and effective compared to the state-of-the-art approach.

## 5 Conclusions and Future Work

The automated and smart biometric-based attendance framework is helpful in increasing the accuracy and speed in a real-time environment. We have employed face detection and recognition in the proposed framework to take the attendance

of students. The proposed system works adequately in different facial expression scenarios. There is still scope for further advancement for recognizing the face of student in the presence of varying poses, illumination variation, occlusion, etc., in the classroom. In the future, this method can be used in corporate companies in terms of various services such as staff security and staff attendance. More advancements can be done in the proposed approach in such a way that after the attendance is recorded, a quick notification would be popped up on the user's cell phone about their presence status and if not notified, he or she can immediately inform the teacher about it. Night vision can be added to reduce the errors in recognizing the faces.

## References

1. Epstein & Sheldon (2002) Present and accounted for: improving student attendance through family. *J Edu Res* 95(5):308–318
2. Ready D (2010) Socioeconomic disadvantage, school atte, early cognitive development: the differential effects of school exposure. *Soc Edu* 83(4):271–286
3. Bruner C, Discher A, Chang H (2011) Chronic elementary absenteeism: a problem hidden in plain sight. Attendance Works and Child & Family Policy Center
4. Kambi BIL, Chunsheng G (2017) Enhancing face identification using LBP and K-nearest neighbors. *J Imag* 3(37):1–12
5. Jain SK, Joshi U, Sharma BK (2011) Attendance management system. Masters Project Report, Rajasthan Technical University, Kota
6. Bhalla V, Singla T, Gahlot A, Gupta V (2013) Bluetooth based attendance management system. *Int J Innov Eng (IJET)* 3(1):227–233
7. Bah SM, Ming F (2020) An improved face recognition algorithm and its application in attendance management system. *Array* 5:1–7
8. Joardar S, Chatterjee A, Rakshit A (2015) A real-time palm dorsa subcutaneous vein pattern recognition system using collaborative representation-based classification. *IEEE Trans Instrum Meas* 64(4):959–966
9. Jain K, Ross A, Nandakumar K (2011) Introduction to biometrics. Springer, pp 1–208
10. Jain K, Flynn P, Ross AA (2010) Handbook of biometrics. Springer, New York
11. Schroff, Kalenichenko D, Philbin J (2015) FaceNet: a unified embedding for face recognition and clustering. In: 2015 IEEE on CVPR, Boston, MA, pp 815–823
12. Kanti J, Papola A (2014) Smart attendance using face recognition. *Int J Adv Res Comput Commun Eng* 3(6):7321–7324
13. Mohamed K, Raghu C (2012) Fingerprint attendance system for classroom needs. In: India conference, annual IEEE, pp 433–438
14. Lim T, Sim S, Mansor M (2009) RFID based attendance system. In: 2009 IEEE symposium on industrial electronics and applications, ISIEA, vol 2, pp 778–782
15. Kadry S, Smaili K (2007) A design and implementation of a wireless iris recognition attendance management system. *Info Tech Control* 36(3):323–329
16. Viola & Jones (2004) Robust real-time face detection. *IJCV* 57(2):137–154
17. Jin H, Liu Q, Lu H, Tong X (2004) Face detection using improved LBP under Bayesian framework. In: Proceedings of international conference on image and graphics, pp 306–309
18. Reddy AM, Krishna V, Sumalatha L (2018) Face recognition based on cross diagonal complete motif matrix. *Int J Image Grap Sig Proc* 3:59–66
19. Hoang A, Sriprasertsuk P, Kameyama W (2013) Fish detection by LBP cascade classifier with optimized processing pipeline. *IPSJ SIG Techn Rep* 82(9):1–20

# Automated Medical Symptoms Summarization and Patient Profiling System



Sakshi Gehani and Siba Panda

**Abstract** This article talks about an interactive intelligent agent that was built using the techniques of natural language processing (NLP) and computer vision (CV). Called the DocAssist, this rule-based AI chatbot is a virtual assistant that tries to imitate the doctors' physician assistant. It is an NLP tool for Automated Medical Symptoms Summarization and Patient Logging System, that would help medical practitioners save some valuable time and energy going through a patient's healthcare portfolio. It functions by asking for the required details and accepting speech inputs from the patients. The bot hence collects all the information such as name, age, gender, symptoms being experienced, and previous medical history, and then by using different techniques of NLP, namely regular expressions (Regex) analysis, FastAi text classification, entity extraction, punctuating, and noun chunking, it tries to extract only the useful information (keywords and entities). All the details are then curated and stored in a text file for the doctors quick perusal. Moreover, the patients also get an option to upload any supporting images (pictures or scanned documents), the corresponding super-resolution form of which is later sent to the doctor.

**Keywords** Natural language processing · Physician assistant · Computer vision · FastAi text classifier · Speech recognition · Virtual assistant

## 1 Introduction

The main idea behind this project was to build an NLP tool for Automated Medical Symptoms Summarization and Patient Logging System, that would help medical practitioners save some valuable time and energy going through a patient's healthcare portfolio.

---

S. Gehani (✉)

Narsee Monjee Institute of Management Studies, Mumbai, India

e-mail: [sakshi.gehani1721@nmims.edu.in](mailto:sakshi.gehani1721@nmims.edu.in)

S. Panda

Mentor, Narsee Monjee Institute of Management Studies, Mumbai, India



The program is aimed at conducting a medical interview of the patient before he meets the doctor, insights from which are curated and sent to the doctor's screen for a quick perusal. The bot tries to capture the patient's demographic data—including his name, age, gender, marital status, height, weight, problems/symptoms being experienced, time since the symptoms showed up, current medication (if any), known allergies, known chronic ailments (such as diabetes, hyperthyroidism, hypertension, hormonal imbalances), and medical history (previous diagnosis, surgeries, etc.).

The patient is allowed to dictate the answers to the bot's queries. The program then converts these recorded audio files to text using Python's Speech Recognition wrapper library (primarily using Google's speech recognition API). After this speech-to-text translation, the bot automatically extracts only the useful information/keywords from the text data generated. For this, it uses several different natural language processing techniques such as entity extraction, noun chunking, text summarization using FastAi classifier, and Regex operations, depending upon the type of information.

It then creates a medical portfolio for the patient with the information extracted in the previous steps. Moreover, the patient is allowed to upload any pictures (especially in case of skin and eye disorders) and/or any scanned documents, including any previous hospital discharge slips, prescriptions from other doctors, medical diagnosis (/laboratory) reports, etc. through a Flask-enabled Web Application page. The uploaded images are further passed through a GAN-based network (SRGAN) that enhances the resolution of the image (in case it is blurred or of low quality/resolution) so that the doctor can have a proper inspection of the details provided.

## 2 Motivation and Goals

### 2.1 *The Problem*

It is generally observed that a medical practitioner having a huge client base usually finds it difficult to function without at least 2–3 physician assistants.

This is majorly due to the fact that most of the patients visiting him for consultations tend to explain their symptoms and problems extensively and in a very descriptive manner in order to avoid having any vital detail being missed out on. It is seen that instead of describing the issue in a word or two, they usually choose to explain every little symptom in detail, e.g., If a patient is suffering from loss of appetite (medically termed as “anorexia”), he is quite likely to express this issue by saying that he has not been feeling hungry or not having the urge to eat. This is more so because the patients are not familiar with the medical jargons and neither do they understand the value of the doctor's time.

Thus having an AI bot that imitates a physician assistant not only saves the doctor some valuable time but also the effort of repeatedly asking the same set of questions from each individual patient. It is also essential for the doctor to maintain

his sanity and calm when the patient is too talkative. Apart from saving the doctor some pennies (that would otherwise be spent in employing the physician assistants), having this system implemented helps the doctor get everything (the patient's profile and supported reports/documents) right at his desk before the patient arrives in his chamber.

Moreover, to mitigate the problem that arises from the patients not having a good-quality camera to capture the required pictures and/or scan the reports, or clicking and up-loading blurred images, the system also supports an image super-resolution feature that aims at enhancing the image quality before it is passed on to the doctor.

## 2.2 Objectives

The main objectives of the project were as follows.

- To build an AI bot that mimics a physician assistant.
- To conduct medical interrogations of patients and extract useful information from what they dictate, by using various NLP techniques.
- To summarize the patients' medical history and symptoms being experienced without missing out any detail or using unnecessary words. This is done in two ways: (i) using entity extraction and noun chunking and (ii) using FastAi text classifier trained on a medical symptoms dataset to categorize the most common symptoms from the various definitions that people use to describe them.
- To allow the patients to be able to send/upload (in high quality) any supporting/required images and/or scanned documents including previous consultations/prescriptions, pathological diagnosis reports, hospital discharge slips, surgery reports, etc., from the comfort of their home.

## 2.3 Prospective Client(s)

This kind of an Automated Patient Profiling System is targeted toward being used by medical practitioners:

- Who have a huge client-base, and who want to save some time and energy by not having to repeatedly interrogate and summarize every patient's problems/cases before understanding the main problem and prescribing some treatment consequently.
- Who take virtual consultations where there is no human assistant to help them with the above-mentioned task. In the absence of an assistant, online appointments become much more time consuming pertaining to the delays due to technical glitches, Internet connectivity issues, etc.

- Who have patients visiting them from afar (different countries/states/cities). In such cases where the conditions are not so favorable for the patient to visit physically, the doctor can have him visit a Web site that automatically curates all his details and sends the generated profile and supported documents to the doctor. The doctor can then simply and quickly review his profile and decide if a physical visit is urgently needed or not.

## ***2.4 Applications/Use-Cases***

- In telemedicine—doctor already has all the necessary information before the patient actually joins the call.
- Cases where the patient cannot physically visit the doctor (in cases of say some mobility restrictions or a pandemic)
- Cases where the patient wants to simply take a second opinion.
- Cases where either the patient (suffering from a chronic ailment or life-threatening disorder like cancer) or the doctor is home quarantined (looking at the present scenario of the COVID-19 pandemic)
- Cases where the patient lives afar and can only afford a visit in case of an urgent need for physical check-up.
- For international patients—to confirm whether or not the surgery machinery is available so that they can plan their travel accordingly, to consult and take alternate therapy (Ayurvedic treatment/Yoga) without having to worry about visa issues.

## **3 The Solution**

### ***3.1 Python Packages Used***

- ffmpeg
- scipy
- SpeechRecognition Wrapper—Google Web Speech API
- Spacy
- Punctuator
- FastAi
- Scispacy
- Matplotlib
- PIL
- ISR
- flask

3.2 Tasks Involved

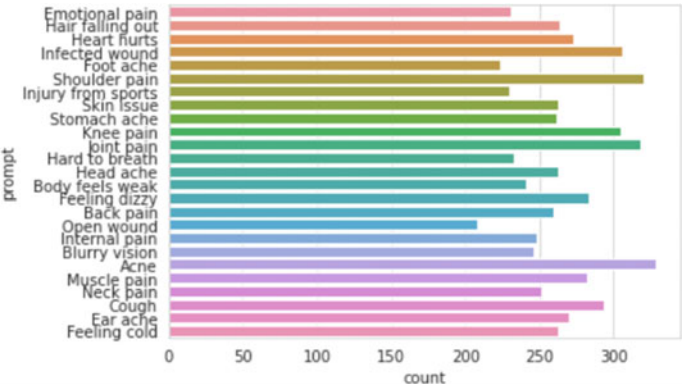
Building a solution for the aforementioned problem involved the following tasks:

- 1. Defining utility functions—that provide GUI for audio input, to accept audio input and save it as a .wav file for future reference, and to convert the speech into text. For these functions, the Python libraries used were ffmpeg, scipy, and SpeechRecognition.
- 2. Defining utility functions—for extracting necessary information from the text generated, including name, age, gender, weight, height, marital status, etc. For this we used NLP techniques such as entity extraction, noun chunking, Regex operations, and conflict-resolution dictionaries.
- 3. Defining functions to summarize/extract medical symptoms—(i) using Python’s Punctuator package, entity extraction, and noun chunking and (ii) using FastAi text classifier trained on a medical symptoms dataset to classify text descriptions of medical symptoms according to the category of the ailment being described. The dataset was obtained from an open-source data repository called “appen” and for the processing one of the spacy pretrained models, specifically trained on (scientific) medical documents—“en\_core\_sci\_sm,” was used (using Python’s Scispacy package). (see Figs. 1, 2, and 3).
- 4. Defining a caller function to sequentially call all the utility functions and to capture all the details of the patient’s profile, that are finally written onto a text file.
- 5. Creating APIs for image super resolution and flask deployment. For implementing image super-resolution, the RRDB model (with pretrained weights) of ISR library has been used.

	file_name	phrase	prompt	overall_quality_of_the_audio	speaker_id
0	1249120_43453425_58166571.wav	When I remember her I feel down	Emotional pain	3.33	43453425
1	1249120_43719934_43347848.wav	When I carry heavy things I feel like breaking ...	Hair falling out	3.33	43719934
2	1249120_43719934_53187202.wav	there is too much pain when i move my arm	Heart hurts	3.33	43719934
3	1249120_31349958_55816195.wav	My son had his lip pierced and it is swollen a ...	Infected wound	3.33	31349958
4	1249120_43719934_82524191.wav	My muscles in my lower back are aching	Infected wound	4.67	43719934
...	...	...	...	...	...
6656	1249120_14353703_45949288.wav	I feel a burning sensation in my guts about 2 ...	Stomach ache	3.80	14353703
6657	1249120_15004831_26452554.wav	I have a split on my thumb that will not heal.	Open wound	3.80	15004831
6658	1249120_15004831_64958100.wav	I feel a lot of pain in the joints.	Joint pain	3.80	15004831
6659	1249120_15830408_92962528.wav	The area around my heart doesn't feel good.	Heart hurts	3.80	15830408
6660	1249120_15004831_80093096.wav	I complain alot with skin allergy	Skin issue	4.00	15004831

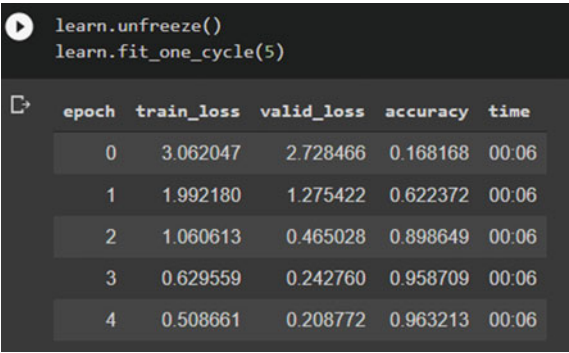
6661 rows x 5 columns

Fig. 1 Dataset (taking only the required columns)



**Fig. 2** Bar Plot—ailment categories versus frequency

**Fig. 3** Fine-tuning the fastai.text\_classifier\_learner() on the symptoms description dataset



3.3 Models/Algorithms Used

Primarily four different models were used:

- i. *Spacy NLP model*—for performing tokenization, pos-tagging, lemmatization, noun chunking, and named entity recognition.
- ii. *FastAi text classifier*—to summarize medical symptoms by classifying text descriptions according to the category of ailment being described.
- iii. *Punctuator model*—a pretrained bidirectional recurrent neural network to add punctuations to unsupervised data.
- iv. *Super-resolution generative adversarial network (SRGAN)*—for upscaling images such that there is negligible information loss, and improving the image quality by converting low-resolution images to higher resolution.

### 3.4 Techniques Used

The foremost technique used was speech-to-text conversion, also known as **speech recognition**. Speech recognition is an interdisciplinary subfield of computer science and computational linguistics that develops methodologies and technologies that enable the recognition and translation of spoken language into text by computers. Most modern speech recognition systems rely on what is known as a hidden Markov model (HMM). This approach works on the assumption that a speech signal, when viewed on a short enough timescale (say, ten milliseconds), can be reasonably approximated as a stationary process. In a typical HMM, the speech signal is divided into 10 ms fragments. The power spectrum of each fragment is mapped to a vector of real numbers known as cepstral coefficients. The final output of the HMM is a sequence of these vectors. Lastly, to decode the speech into text, groups of vectors are matched to one or more phonemes—a fundamental unit of speech.

Another major technique used is **text classification**. Text classification also known as text tagging or text categorization is the process of categorizing text into organized groups. By using natural language processing (NLP), text classifiers can automatically analyse text and then assign a set of predefined tags or categories based on its content.

Lastly the technique of **image super-resolution** was used. The estimation of a high-resolution (HR) image from a single low-resolution (LR) counterpart is referred to as super-resolution (SR). LR is a single image input, HR is the ground truth, and SR is the predicted high-resolution image. When applying ML/DL solutions, the LR images are generally the down-sampled HR image with some blurring and noise added to them. The different methods for image super-resolution that have evolved over time are—(i) interpolation method, (ii) SRCNN—an application of faster CNN in super-resolution, (iii) SRResNet (CNN with skip connections) and subpixel convolution for unsampling, and (iv) SRGAN. Here, we have used the SRGAN model. Super-resolution generative adversarial network [SRGAN], like any typical GAN model, consists of two components—a generator that learns to generate super-resolution (SR) images from low-resolution (LR) images as close as possible to the high-resolution (HR) counterparts and a discriminator that learns to distinguish the generated SR images from the real (input) images. SRGAN model also combines perceptual loss [please refer to Eq. (1)] with generative or adversarial loss [please refer to Eq. (2)] for the computation of its loss [please refer to Eq. (3)].

$$l_{\text{VGG}/i,j}^{\text{SR}} = \frac{1}{w_{ij}H_{ij}} \sum_{x=1}^w \sum_{y=1}^H \left( \phi_{ij}(I^{\text{HR}})_{xy} - \phi_j(G_{\theta}(I^{\text{LR}}))_{xy} \right)^2 \quad (1)$$

$$l_{\text{Gen}}^{\text{SR}} = \sum_{n=1}^N -\log_{\theta_D}^D(G_{\theta_D}(I^{\text{LR}})) \quad (2)$$

$$l^{\text{SR}} = l_x^{\text{SR}} + 10^{-3} l_{\text{Gen}}^{\text{SR}} \quad (3)$$

## 4 Future Aspirations

- (a) To provide Hindi-language support and support for various other regional languages also so that there is no language barrier problem.
- (b) Website deployment with a database and logging system to store, retrieve, and manage patients' medical portfolios.
- (c) To make available different versions for different medical specializations.
- (d) Deployment on any telemedicine portal/app with chat support.

## 5 Conclusion

The solution suggested in this paper would thus help medical practitioners with summarizing the patients' symptoms and creating their medical profiles, i.e., by automating the work generally done by human assistants. This not only saves the doctor some valuable time, energy, and money (no hiring cost) but also helps him deal with the patients in a more efficient manner.

**Acknowledgements** The authors wish to thank Prof. Sarada Samantaray (Associate Dean and HOD, Data Science, Narsee Monjee Institute of Management Studies) for his guidance and for providing the necessary facilities for the project.

## References

1. Barr PJ et al (2017) Sharing annotated audio recordings of clinic visits with patients-development of the open recording automated logging system (ORALS): study protocol. *JMIR Res Protocols* 6(7):e121. <https://doi.org/10.2196/resprot.7735>
2. Ledig C (2016) Photo-realistic single image super-resolution using a generative adversarial network. *arXiv e-prints*
3. [https://www.cdc.gov/nchs/ahcd/datasets\\_documentation\\_related.htm](https://www.cdc.gov/nchs/ahcd/datasets_documentation_related.htm)
4. <https://indiaai.gov.in/article/how-to-convert-speech-to-text>
5. <https://browse.welch.jhmi.edu/datasets/ehr-databases>
6. <https://medium.com/curai-tech/nlp-healthcare-understanding-the-language-of-medicine-e9917bbf49e7>
7. <https://github.com/IBM/watson-voice-bot>
8. <https://appen.com/open-source-datasets/>
9. <https://nlp.fast.ai/>
10. <https://spacy.io>

# Smart Shoe Provides Vision to Visionless Person



Roshani K. Dharme, Jaya R. Surywanshi, Harsha C. Kunwar,  
and Yogesh H. Palve

**Abstract** Smart shoe is a system which makes life of blind person independent and easy. The paper propose a system that helps blind person and it provides vision to visionless people. The paper includes six modules: First module provides GPS navigation assistance to blind person with use of voice command, and if the person gets lost, we can track that person's location with the help of shoes. Second module is design for obstacle detection, obstacle may be potholes, walls vehicles, etc. and when the obstacle is detected the blind person will be notified by the voice message. Third module is design to work for fire and water detection, where sensor will detect fire and water and the blind person will get informed through voice message. Fourth module is panic touch, when person is in uncomfortable situation or he/she falls somewhere then person will click five times the power button then the current location will be send to his close ones. Fifth module provides sensor which convert's natural energy into electrical energy. Son by using this footsteps power generation system, we are generating and storing electricity. Sixth module provides information about travelled footsteps, distance, calories burned, person's weight, total fuel saved and if shoes are missing we can find it with the help of alarming system and last with the help voice commands we will inform the blind person at regular time period what the time is and whether it is day or night. Smart shoe is totally based on voice commands. It is portable, easy to use, user friendly and cheap.

**Keywords** Piezoelectric · Pothole · IR sensor · Step tracker

---

R. K. Dharme (✉) · H. C. Kunwar  
G.H. Rasoni Institute of Engineering and Technology, Nagpur, India  
e-mail: [roshani.dharme@raisoni.net](mailto:roshani.dharme@raisoni.net)

H. C. Kunwar  
e-mail: [harsha.kunwar@raisoni.net](mailto:harsha.kunwar@raisoni.net)

J. R. Surywanshi · Y. H. Palve  
MVPS's KBT COE, Nashik, India  
e-mail: [suryavanshi.jaya@kbtcoe.org](mailto:suryavanshi.jaya@kbtcoe.org)

Y. H. Palve  
e-mail: [palve.yogesh@kbtcoe.org](mailto:palve.yogesh@kbtcoe.org)



# 1 Introduction

Presently, blind people use a stick as a tool for directing them when they move for a walk. Here, we developed shoes which can be proved to be more efficient and helpful than the conventional one. The primary goal of this paper is to provide an eye to blind person by delivering safe fearless and independent walk, tracking the health of the person to get a healthier and more active life by just wearing the shoe and also it deals with the problems faced by them to walk like the normal human beings. It is a way to give an eye to the blind people with the aid of technology.

**The first feature is generating electricity** while walking, in which we are using piezoelectric plates that will generate electricity when the pressure is applied on it by the foot, it uses piezoelectric effects. It converts electrical energy into natural energy. It can charge all the sensors and also mobile phone on the way. Piezoelectricity is the electric charge that amasses in certain strong materials because of applied mechanical pressure. The word piezoelectricity implies power coming about because of pressing factor and inactive warmth.

**The Second feature is GPS navigation and tracking** in which by adding GPS and GSM we can easily track the person location so it is very helpful to find the lost person just with the help of shoes and it will be easy to blind person to reach their destination with voice map. GPS means "Global Positioning System". GPS is a satellite route framework used to decide the ground position of an article. GPS innovation was first utilized by the United States military during the 1960s and ventured into regular citizen use throughout the following not many years. Today, GPS beneficiaries are numerous for some business items, like cars, cell phones, practice watches, and GIS gadgets.

GSM stands for Global System for Mobile Communications, one of the main computerized cell frameworks. GSM utilizes tight band TDMA, which permits eight concurrent approaches a similar radio recurrence. GSM was first present in 1991. As of the finish of 1997, GSM administration was accessible in excess of 100 nations.

**The Third element** is snag location; it can improve the versatility just as security of outwardly hindered individuals particularly in new climate. At the point, when the snag is identified and restricted and the data of hindrances will be shipped off outwardly disabled individuals by utilizing voice order, to identify the obstruction we are utilizing Ultrasonic sensors. As the name shows, ultrasonic sensors measure distance by utilizing ultrasonic waves. The sensor head produces an ultrasonic wave and gets the wave reflected back structure the objective. Ultrasonic sensors measure the distance to the objective by estimating the time between the outflow and gathering.

**The fourth component** is water location. In water discovery, if the water were to get feeling of the water then the water wire will offer yield to the visually impaired individual and voice. Soil dampness sensors measure the volumetric water content in soil. Since the direct gravimetric estimations of free soil dampness requires eliminating, drying, and weighting of an example, soil dampness sensors measure the

volumetric water content by implication by utilizing some other property of the dirt, for example, electrical obstruction dielectric consistent, or cooperation with neutrons, as an intermediary for the dampness content.

**The fifth component** is fire identification; in fire discovery, we are utilizing temperature sensor that identify the temperature and fire from a particular source and adjust the gathered data will be informed to the visually impaired individual. Temperature sensor is a gadget which is planned explicitly to gauge the hotness of chilliness of an article. LM35 is a PRECISION, for example, temperature sensor with its yield relative to the temperature (in  $^{\circ}\text{C}$ ). With LM35, the temperature can be estimated more precisely than with a thermistor.

**The sixth element** is pothole recognition in which the potholes are distinguished and daze individual will advise about the pothole. The TSOP 1738 is an individual from IR controller collector arrangement. This IR sensor module comprises of a PPIN diode and a pre-speaker which are installed into a solitary bundle. The yield of TSOP is dynamic low, and it gives  $+5\text{ V}$  in of state. At the point when IR waves, from a source, with a middle recurrence of  $38\text{ kHz}$  episode on it, its yield go low, lights coming from daylight, rich lights and so on may make aggravation it and result in bothersome yield in any event, when the source isn't communicating IR signals. A band pass channel, an integrator stage and a programmed acquire control are utilized to smother such aggravations.

**The seventh feature is day/night and time announcement** in which in every hour or manually by voice command the announcement will generate in blind person's ear to tell the current time and will disclose whether there is day or night, by this feature the blind person will updated.

**The eight feature is health tracking** in which it will count number of steps blind person taken and calculate burned calories and distance travelled with one-week data graph, which will be displayed in the android application. To count the number of steps travelled, we are using steps tracker switch, and to find the burned calories and distance travelled, we are using different formulas.

**The ninth feature is panic touch** in which when the blind person feels uncomfortable or sees himself/herself in panic situation or the blind person falls somewhere then by touching mobile power button five times then the emergency message will be sent to the close ones. This all highlights primarily associated with Brain of our undertaking called Arduino Nano. Arduino Nano is a little, viable, adaptable and breadboard amicable Microcontroller board dependent on ATmega328p (Arduino Nano V3.x)/at mega 168 (Arduinio Nano V3.x). It accompanies a working voltage of  $5\text{ V}$ , be that as it may, the info voltage can differ from  $7$  to  $12\text{ V}$ . Advanced shoes, for the most part centres around giving individuals to walk unquestionably and autonomously by wearing a couple of shoe and to get a better and more dynamic life. The individual who will wear the shoe will get tell by Bluetooth and the android application will be an arbiter among shoe and the visually impaired individual.

## 2 Related Works

A huge works have been accomplished for daze individuals, so that visually impaired individual can undoubtedly move one spot to another without being weight of others. In any case, there is no finished work together. It implies that, on the off chance that one work is accomplished for GPS route other is for snag location. Along these lines, the work is discrete not coordinated [1].

Item Detection System for blind People, various sicknesses brought about by visual disability and visual impairment has been colossally decreased; however, there are numerous individuals who are in danger old enough related visual impedence. Visual data is the reason for most navigational assignments, so important data about the general climate isn't accessible to outwardly weakened individuals is at weakness [2].

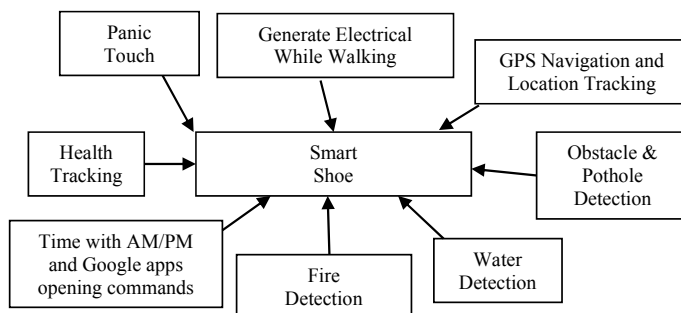
Bite (2012) proposed the shrewd white stick, called blind recognize that joins GPS innovation, interpersonal interaction and supersonic sensors to assist outwardly disabled individuals with exploring public spaces [3].

Wahab et al. (2014) built up a stick could speak with clients through voice alarm and vibration signal. Ultrasonic sensors are utilized to distinguish deterrent in front, since ultrasonic sensor are acceptable in recognizing hindrance In couple of metres reach and this data will be sent as voice signal. This voice signal is send by means of speaker to the client [4].

**Produce power and putting away.** Another methodology, which wipes out the force wiring issue, is creating and putting away electric energy at the actual gadgets by searching waste energy structure According to assessments, 67 watts of force are accessible in the heel development of a normal (68 kg) individual strolling at an energetic speed, yet a couple of rates of their energy is appropriate for the sustenance of an electronic gadget. This issue is drawn closer by utilizing the energy from the weight move during a stage to perform helpful work. So this paper is centred around the advancement of a “simple to utilize” and “modest universally useful” gadget for the capacity administration of the gathered energy with specific applications on shoes [5].

Gangwar (2011) designed a shrewd stick for dazzle which can give early admonition of a snag utilizing infrared (IR) sensors. Subsequent to recognizing the hindrances, the stick cautions the outwardly hindered individuals utilizing vibration signals. Nonetheless, the brilliant stick zeroed in just for deterrent location but it isn't helping for crisis purposes required by the visually impaired. And furthermore, the IR sensor is not actually effective enough since it can recognize just the closest deterrent in brief distance [6].

Benjamin (2011) had built up a shrewd stick utilizing laser sensors to identify the snags and down controls. Obstruction identification was signalized by a high pitch “Blare” utilizing a receiver. The plan of the laser stick is basic and natural. The stick can just distinguish hindrance, yet can't offer intellectual and mental help. There exists just signal sound that is triggers and snag there is no any help to coordinate the [7].



**Fig. 1** Flow diagram of smart shoe

### 3 Methodology

Piezoelectric plates are used to generate the electricity by applying pressure using foot. GPS navigation and tracking has been added to track the person location so it is very helpful to find the lost person just with the help of shoes, and it will be easy to blind person to reach their destination with voice map. Impediment recognition is finished utilizing ultrasonic sensors. Soil dampness sensor is utilized to gauge the volumetric water content. Temperature sensor used to detects the temperature and fire from a specific source. IR sensor is used to detect potholes. Steps tracker switch is used to find the burned calories and distance travelled. In addition time, announcement has been given to blind person on hour's basis. If blind person feels uncomfortable, then he/she touches power button 5 times to generate emergency message to contact who can help him/her.

Every one of these highlights essentially associated with Arduino Nano. Arduino Nano is a little, viable, adaptable and breadboard cordial Microcontroller board dependent on ATmega328p (Arduino Nano V3.x)/AT mega 168 (Ardunio Nano V3.x). It accompanies a working voltage of 5 V, notwithstanding, the information voltage can differ from 7 to 12 V (Fig. 1 and Table 1).

#### Output

See Figs. 2 and 3.

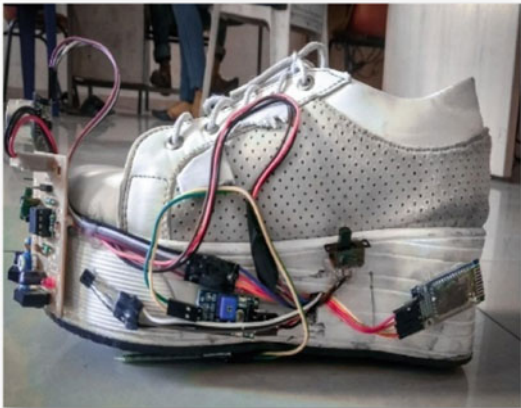
### 4 Conclusion

Smart shoe provides an eye to blind person by delivering safe fearless and independent walk, tracking the health of the person to get a healthier and more active life by just wearing the shoe and also it deals with the problems faced by them to walk like the normal human beings. It is a way to give an eye to the blind people with the aid of technology. Smart shoe is tested successfully. It is ready to overcome all

**Table 1** Test cases and debugging

Test case ID	Test objective	Input data	Expected result	Actual result	Status
TC_01	Check Obstacle Sensor	Placed object or obstacle like wall, human, etc.,	Obstacle is here will speak out when obstacle is detected	Obstacle is detected	PASS
TC_02	Check Pot Hole Sensor	Keep the sensor near any hole	Pot hole is here will speak out when pot hole is detected	Pot hole is detected	PASS
TC_03	Check Fire Sensor	Keep the sensor near fire	Fire is here will speak out when fire is detected	Fire is detected	PASS
TC_04	Check Water Sensor	Keep the sensor near water	Water is here will speak out when water is detected	Water is detected	PASS
TC_05	Check Piezoelectric Element	Applied pressure on the Piezoelectric Element	Electricity will generate	Electricity Generated	PASS
TC_06	Check Health Tracker Switch	Press the switch and release	Step will counted	Step is counted	PASS
TC_07	Check QR Code	Scan the QR code	Information will be displayed	Information is displayed	PASS

**Fig. 2** Smart shoe side view



**Fig. 3** Smart shoe front view



the problems, which is face by blind people. It has a great commercial potential and business plan. Smart shoes can be a great aid to these people, and it can make a great impact on the country and help us walk towards a better and bright future.

## References

1. (2017) Int J Eng Sci Comput
2. Udgirkar S, Sarokar S, Gore S, Kakuste D, Chaskar S (2016) Object detection system for blind people. Int J Innov Res Comput Commun Eng (An ISO 3297: 2007 Certified Organization) 4(9)
3. Chew S (2012) Proposed the smart white cane, called Blind spot that combines GPS technology, "Electronic Path Guidance for Visually Impaired People". Int J Eng Sci (IJES) 2(4):9–12
4. Wahab MH, Talibetal AA (2013) A review on an obstacle detection in navigation of visually impaired. Int Org Sci Res J Eng (IOSRJEN) 3(1):01–06
5. ISO 3297:2007 Certified (2017) Int J Adv Res Comput Commun Eng 7(11)
6. Gangwar S (2013) Designed a smart stick for blind which can give early warning of an obstacle using infrared (IR) sensors, "A smart infrared microcontroller-based blind guidance system". Hindawi Trans Active Passive Electron Compon 3(2):1–7
7. Benjamin, Shimi SL, Chatterji S (2014) Design of microcontroller based Virtual Eye for the Blind. Int J Sci Res Eng Technol (IJSRET) 3(8):1137–1142

# Relevance of Disruptive Technologies Led Knowledge Management System and Practices for MSME



Ravi Shankar Jha , Priti Ranjan Sahoo , and Smrutirekha

**Abstract** In today's dynamic and hyper-competitive era, knowledge is the backbone of an organization and provides unflagging support in inclusive growth including financial and non-financial terms to the organization. It is essential for organizations including MSME firms to cultivate and nurture knowledge-enriching system within organizations to extract knowledge from external stakeholders, internal stakeholders (like employees), and customers through organization learning ecosystem to understand the essence of past, administer the present, and envision the future roadmap. In recent times, knowledge management practice garners significant attention at leadership level and being considered as strategic initiative in organization which helps in achieving competitive advantage and propel inclusive growth. It also highlights on various factors which are being positively influenced by adopting knowledge management practices in day-to-day operations such as organization's performance, innovation capability, improve products and services, and job satisfaction. Therefore, organization needs to define a standard practice to acquire, retain, and leverage the knowledge through "knowledge management" initiative constituting creation, acquisition, transfer, storage, and sharing of tacit and external knowledge to reap the benefits across business value chain. To promote the growth of knowledge management in organization's setting, knowledge management system (KMS) and practice play paramount role with the support of disruptive technologies. Additionally, disruptive technologies arm not only simplify activities to manage knowledge but also enhance experience of the business users. Ubiquitous penetration of disruptive technologies, primarily accelerated by rapid adoption of digital channels (accelerated through COVID), advent of FinTech and InsurTech and adjacencies that many players are getting into with their scale. The small, medium, and micro-sized firms must explore digital technologies to improve customer experience, time to market and reduce costs by honing organizational knowledge. Also, they had to find additional dollars to invest in Next-Gen Technologies-led platforms including Cloud, API, Agile, DevOps, AI/ML, intelligent automation to stay relevant for their customers.

---

R. S. Jha (✉) · P. R. Sahoo · Smrutirekha

KIIT School of Management, KIIT University (Institution of Eminence), Bhubaneswar, India

P. R. Sahoo

e-mail: [prsahoo@ksom.ac.in](mailto:prsahoo@ksom.ac.in)

Precisely in MSME context, knowledge management is one of the most important factors to secure sustainable growth, create unique offerings (products and services), accelerate business performance and attain inclusive growth. Whether MSME firms are ready to take up the challenge to marry existing knowledge management system and practice with new age disruptive technologies to secure first mover advantage and set an example to Industry and nation for radical growth?

**Keywords** Knowledge management · Disruptive technologies · Machine learning · Micro-small · Medium-sized enterprise (MSME)

## **1 Overview on Knowledge Management, Disruptive Technologies and MSME**

### ***1.1 Knowledge Management***

Knowledge management is pivotal for any organization in the age of dynamic technology, business landscape, customer's preferences, and globalization to name a few. Present day economy is better known as "knowledge economy" which focuses on strategic asset, i.e., "knowledge" over labor or other natural resources. Knowledge provides a good leverage; with the help of this, organizations have competitive edge over its competitors. It provides better returns so the objective of having in knowledge management system is to see that how an organization can make use of knowledge to achieve competitive advantage to get better returns, efficiency, and productivity [10]. The emphasis on knowledge management is also beautifully put by [13, 14], successful organizations are those that consistently create new knowledge, disseminate it widely throughout the organization, and quickly leverage it in new technologies and products.

### ***1.2 Disruptive Technologies***

Technology in essential pillar for successful adoption of KM in Organizational [4, 22, 23]. Knowledge management system (KMS) is building block to setup knowledge management practices and process to bring theories into actions. Most of the MSME organizations are profoundly reckoning on either developing or grooming KMS system or KM tools/framework by exploiting ICT infrastructure and technical capabilities to seamlessly execute knowledge management practices [2]. Moreover, in digital and technology-driven era, future is loaded with Google Home voicebot, Chabot, personalized smart videos for product information, augmented reality product awareness, omnichannel-based solution framework to launch new disruptive products rapidly into the market, blockchain-based solutions for functionalities like claims handling process, subrogation and KYC, and many more. Lastly,



big data, advanced analytics, IoT, and other disruptive technologies helping MSME firms to create smarter environment for creating and managing organizational knowledge on the move in data-driven ecosystem to improve firm's innovation capability with ease in decision making to delight all possible stakeholders at every touch points in the business value chain [8, 9].

### ***1.3 MSME and Notable KM Influencers***

The micro-, small, and medium enterprises (i.e., MSME) serves as backbone for developing and underdeveloped economies. It charts out a multidisciplinary development in terms of social-economic growth of the nation by promoting entrepreneurship, skill upgradation, and technology upgradation and provides line of credit, infrastructure development and modernization leading to scores employment opportunities across sectors. Apparently, MSMEs are integral to supply chain for products and services and avenues for large-scale employment. In the era of globalized knowledge-driven economy, which is prevailed by frequent disruptive changes, MSME firms need to play greater role by extracting data, information, and knowledge from domestic and foreign sources to make notable participation in the global markets, promote innovation, fuel technical capabilities and propel functional and operational performance [6].

In addition to classification of MSME firms in India, it is equally important for government and industrial bodies to create conducive environment to accelerate overall capabilities and build new-generation products and services through collective bodies of knowledge. Few notable influencers (investment in building nimble and agile KM framework, unified KM practices and foster co-development and co-innovation working models with industrial bodies) which will certainly accelerate MSME firms to embrace disruptive technologies in its organizational setup to enrich industrial knowledge and contribute holistically to the growth of economy and nation at large.

### ***1.4 Key Drivers for Adoption of Knowledge Management in MSME***

Knowledge management, one of the crucial drivers for MSME, uniquely tackles the increasing problems of knowledge erosion. Following are few key drivers which compels MSME to pursue knowledge management [16] as one of their central business strategies.

**Competition dynamics** Globalization paired with ever-changing technology and business landscape is the biggest challenges. It is imperative for MSMEs to consistently create new knowledge out of the data, and the contextual information they

must position as organizations that offer something different and unique values to the clients and stakeholders.

**Employee Turnover** Attrition/turnover being a global phenomenon which impacts everyone in the value chain including MSME firms, so to retain business firm needs to have well-defined knowledge management practices and seamless execution to avoid human and other external dependency.

**Succession Planning** Needless to say, without proper succession planning organization will not survive beyond certain points. As [24] asserts, once an enterprise's owner dies or decides to quit, the knowledge within the company might become eroded. Further, the leadership change might drive some managers and employees to quit their jobs. Thus, knowledge management (KM) practices play a pivotal role in securing, storing, discovering CXO level knowledge base which enables in proper and effective succession planning.

**Merger and Acquisition (M & A)** As part of growth story, firms explore organic and inorganic ways to meet the changing demands of customers by adopting various strategic, tactical, and operational strategies to remain exist in business. In course of that, firm either acquires firm of similar nature or merge with larger firm to obtain skill, scale, business, people, and process to remain in existence. In this endeavor, knowledge will be paramount to achieve synergy (among firms) and help to fulfill its objectives by leveraging knowledge management practices.

**Disruptive technology landscape** With changing technology era, storing and sharing tacit or implicit knowledge is quintessential to lay the foundation for building formidable knowledge management system augmented by disruptive technologies to support various business capabilities across the value chain [19, 20]. New age knowledge management systems are complemented by Internet of things (IoT), big data, machine learning, artificial intelligence, and many more to create unique value proposition for clients and stakeholder [1]. Also, MSME has been actively pursuing this as it fosters innovation and continuous improvement and offers agility in their business operations.

## 2 Role of Disruptive Technologies in Shaping KM Practices in MSME

Twenty-first century has started making the leap of research and hard work in the space of technologies. The last decade of twentieth century has seen the Internet revolution and the first decade of current century has witnessed the mobile revolution making a huge impact in the way we capture, acquire, store and share knowledge of all sizes of industries in general and MSME in particular. The IoT-driven smart and connected systems are collecting huge amount of data due to digitization across all business strata. The state-of-the-art processors are now capable to process trillions

and trillions of data which are then modeled and remodeled using analytics engines for all the three forms of analysis (description, predictive, and prescriptive) so that the conversion of data to information makes sense to the industry users [3]. Among all these, the data on cloud ensures that we have the information anytime and anywhere.

The above behavior/adoption sounds great. But unfortunately, the abundance of data leading to information proliferation is leading to distraction from consumers' core values, problems, and real-life problems. That is when knowledge workers are significantly investing time and money to curate and build context-driven content which will add value to the consumer and help them to make informed decision better and faster [17]. That is why the researchers, industry leaders, and thinkers are working on machine learning and other disruptive technologies with a hope to simplify the process of knowledge management (discovery and beyond). This is one of the most popular disciplines of artificial intelligence which is helping teams to discover and manage tacit, explicit, and external knowledge more effectively and efficiently.

In auto industry, ML techniques are significantly accelerating the root cause analysis of a problem and speed up the resolution. The anomaly detection algorithms are analyzing vast amount of drivers and systems data of myriad data types, sizes and quantities far more efficiently than any traditional methods available [15].

MSME financial institutions or multi-national banks operating from various locations of the globe are investing on training their machines and models to create a single integrated knowledge management platform for multilingual workforce to support MSMEs. One of the interesting features like pattern recognition is being used to derive important insights which are helping all the offices of banks [18] and reduce the lending toward NPAs.

Clinical trials for Pharma companies cost them huge both in terms of time and money. But machine learning and in conjunction with other advanced technologies like predictive analytics can identify potential candidates for clinical trials and help researchers draw a desired pool from wide variety of data points on doctors' visits, social media, previous trials, etc. to carry out their research swiftly.

In education sector, emerging and established enterprises are moving from brick-and-mortar to click-and-mortar model and taking numerous initiatives to transform the traditional mode of training and talent enablement. As part of that, Delhi-based Safeducate enterprise used artificial intelligence (AI) operated van for developing relevant competencies and framework to enable learners to achieve a sustainable livelihood [5] by making them job-ready, or ignite the spirit to pursue entrepreneurship. Additionally, Safeducate enterprise launched Car-O-Job, which transforms vans into virtual reality (VR)-led experiential centers for students. Through VR, the van provides glimpses of different jobs and related specification cut across different sectors.

The Government of India has launched "CII Tech Saksham," a project focused on supporting MSMEs in India to become tech-enabled. Being one of its own kind of initiative helps bringing together technology providers closer to MSMEs needs and enables the next wave of growth through synergy and harmony [12]. CII Tech Saksham for MSMEs brings together project partners—Dell Technologies India,

HP India, Intel India, Vodafone Idea Ltd, WhatsApp, and Yes Bank to address technological gaps faced by MSMEs in their growth.

All the above examples are hinting at one underlying need that in digital era information is costly affair, and it is essential to secure information through trusted source at appropriate time irrespective of the industry we cater to make informed decision. Apparently, embracing evolutionary technologies like AI, ML, big data, RPA, blockchain, IoT, AR, VR are all complementing each other to enrich knowledge management ecosystem and craft as state-of-the-art knowledge management systems to achieve innovation, improve business performance, offer best-in-breed products/services and attain competitive advantage in MSME and larger enterprise as well.

### 3 Enablers, Benefits, and Challenges

Based on the research work performed by various notable researchers [7, 13, 14] knowledge is defined as “knowledge is a justified personal belief that increases an individual’s capacity to take effective action.” The objective of knowledge management is to support creation, transfer, and application of the knowledge at enterprise level. The increased pace of data proliferation on the back of digital revolution across industries and boundaries, enterprises have begun adopting “knowledge management” applications and practices to reap the benefits of tacit (internal knowledge) and external knowledge. In current era, most of the MSMEs are using new edge technologies enabled KM system/application to retain enterprise knowledge helping in the growth trajectory of the enterprise. In course of implementing effective and efficient KM systems, MSME realize few key benefits and enablers by working in continuous collaboration with impacted stakeholders to create state-of-the-art KM application/system providing best-in-breed offerings supported by new age technology including machine learning, IoT, AI, and many more (Table 1).

### 4 Discussion

In near future, it would be extremely interesting to observe how the rapidly evolving emerging yet disruptive technologies and systems streamlines organizational knowledge and address the cause; role of induced cognitive ability assisting in research areas, play of advanced natural language processing (NLP) in training the workforce anytime anywhere to reap great dividends, role of blockchain in creating healthy and trustful atmosphere by removing tampering vulnerabilities across transactions, role of augmented reality (AR) in reducing the bottleneck in assembly line by enabling workforce to deliver fast and shorten time to market, role of 3D printing technology helping delivering complex design in less time and commercially viable and impact

**Table 1** Enablers, benefits, and challenges/shortcomings

Enablers	Key benefits	Challenges/shortcomings
Organization structure and culture	Promote digitalization of tacit and external knowledge	1. High cost of new age technology led systems 2. Revamp organization change management framework
Leadership support	Increase accessibility of knowledge	1. Lack of trust in technology 2. Conducive environment for knowledge sharing
Information and communication technology (ICT)	Offers competitive advantage	Unavailability of skilled workforce
Technology (Digital Stack—Including Machine learning, AI, IoT, SMAC, etc.)	Increase operational efficiency and agility	Additional investment on training existing workforce
Knowledge academy	Speed Time to Market—Reduce Product and Service life cycle	Financial and non-financial Support from Government, Policymakers, and Industrial bodies
Training catalog and material	Enhance product and service quality and mitigate risk	Upfront Investments to institutionalize training pedagogy
Knowledge management Infrastructure	Reduce people dependency and attrition risk	1. Support from different strata of leadership at Operational, Tactical, and Strategic level 2. Robust governance structure

of robotic process automation (RPA) in reducing human error and enables organizations to leverage workforce and organizational knowledge to focus on core areas. There will be more addition to the list as technology pave its way on shaping organizational knowledge management processes and practices in knowledge-driven economy. Also, it will help in forming digital brain powered by knowledge for enterprising and sustainable tomorrow for MSME firms and others in tangible and intangible ways.

## 5 Conclusion

Many movies have predicted a future in which machines and data will rise and govern every aspect of human life. With recent changes across industries and explosive growth of disruptive technologies, knowledge management practice and processes will be game-changer in shaping MSME enterprise's future [11]. Also, growth of

MSME enterprises is of utmost importance for a growing global economy. Based on various research works, it is evident that growth is sustainable given the fact that knowledge and data flows freely across the organization value chain. Hence, knowledge management practices play a pivotal role in not only creating and acquiring but also storing, transferring, sharing, and disseminating know-how across the value chain and touch every key stakeholder. To make this knowledge management process effective and efficient, technology is one of the pragmatic catalysts in strengthening the knowledge-led ecosystem. In recent times, it has been strongly witness that MSME exploring range of disrupting integrations by converging knowledge management with machine learning and other emerging yet disruptive technologies to sharp and hone “knowledge” in the KM systems and exploit benefits by uplifting barriers to achieve quality and know-how level of products and services conceive and build across the business value chain. With all these, technology-enabled KM helps to offer scale, innovative products or services and agility to sectors such as education, food, manufacturing, and few more key MSME-dominated sectors to remain innovative and profitable. In the coming days, MSME will implement KM practices in structured manner using disruption technologies to foster continuous improvement, innovation, and achieve sustainable competitive advantage and contribute to the growth of nation and economy [5]. Apparently, MSME entrepreneurs have been relentlessly using myriad technologies and domain capabilities to innovate their products and services, thereby giving a boost to their business and help to offer scale and support to sector at large [21]. In a nutshell, economy of a nation prospers only when MSME sectors registered steady growth provided organization learning and culture accelerate focus in retaining and flourishing organizational knowledge by leveraging proven knowledge management process and practices supported by range of disruptive technologies for better and smarter tomorrow.

## References

1. Abagissa AT, Behura A, Pani SK (2018) IoT based smart agricultural device controlling system. In: Proceedings of the international conference on inventive communication and computational technologies, ICICCT 2018, pp 26–30. <https://doi.org/10.1109/ICICCT.2018.8473126>
2. Centobelli P, Cerchione R, Esposito E (2017) Knowledge management systems: the hallmark of SMEs. *Knowl Manag Res Pract* 15(2):294–304
3. Cerchione R, Esposito E, Spadaro MR (2016) A literature review on knowledge management in SMEs. *Knowl Manag Res Pract* 14(2):169–177. <https://doi.org/10.1057/kmrp.2015.12>
4. Chan I, Chao C-K (2008) Knowledge management in small and medium-sized enterprises. *Commun ACM* 51(4):83–88
5. Dasari S, Jigeesh N, Prabhukumar A (2016) Knowledge management practices in different SMEs: an empirical study in India. *Int J Knowl Manag Stud* 7(3–4):167–191. <https://doi.org/10.1504/IJKMS.2016.082340>
6. Gold AH, Malhotra A, Segars AH (2001) Knowledge management: an organizational capabilities perspective. *J Manag Inf Syst* 18(1):185–214
7. Huber GP (1991) Organizational learning: the contributing processes and the literatures. *Organ Sci* 2(1):88–115. <https://doi.org/10.1287/orsc.2.1.88>

8. Jha RS, Sahoo PR (2020) Internet of Things (IoT)—enabler for connecting world. In: ICT for competitive strategies, pp 1–7. <https://doi.org/10.1201/9781003052098-1>
9. Jha RS, Sahoo PR (2021) Influence of big data capabilities in knowledge management—MSMEs. In: ICT systems and sustainability. Springer, pp 513–524
10. Massa S, Testa S (2009) A knowledge management approach to organizational competitive advantage: evidence from the food sector. *Eur Manag J* 27(2):129–141. <https://doi.org/10.1016/j.emj.2008.06.005>
11. Massaro M, Handley K, Bagnoli C, Dumay J (2016) Knowledge management in small and medium enterprises: a structured literature review. *J Knowl Manag* 20(2):258–291. <https://doi.org/10.1108/JKM-08-2015-0320>
12. Ministry of MSME and Confederation of Indian Industry (2021) CII TechSaksham. <http://cii.techsaksham.in/>
13. Nonaka I (1994) A dynamic theory of organizational knowledge creation. *Organ Sci* 5(1):14–37. <https://doi.org/10.1287/orsc.5.1.14>
14. Nonaka I, Byosiene P, Borucki CC, Konno N (1994) Organizational knowledge creation theory: a first comprehensive test. *Int Bus Rev* 3(4):337–351
15. Obaid AMH, Pattnaik PK, Pani SK (2018) Towards on cloud of things: survey, challenges, open research issues, and tools. In: Proceedings of the international conference on inventive communication and computational technologies, ICICCT 2018, pp 1458–1463. <https://doi.org/10.1109/ICICCT.2018.8473212>
16. Roşu SM, Drăgoi G, Guran M (2009) A knowledge management scenario to support knowledge applications development in small and medium enterprises. *Adv Electr Comput Eng* 9(1):8–15. <https://doi.org/10.4316/aecce.2009.01002>
17. Roşu SM, Drăgoi G, Guran M (2009) A knowledge management scenario to support knowledge applications development in small and medium enterprises. *Adv Electr Comput Eng*. <https://doi.org/10.4316/aecce.2009.01002>
18. Roşu SM, Drăgoi G, Roşu L (2010) Virtual enterprise networks solutions to support the virtual teams work. *Proc Manuf Syst* 5(2):77–80
19. Sahoo PR, Lenka SK, Pradhan BB (2017) Transactional and psychological concerns of all parties involved in opaque hotel room distribution. *Int J Econ Res* 14(8):1–9
20. Sahoo PR, Pradhan BB, Lenka SK, Patra SK (2017) Review research on application of information and communication technology in tourism and hospitality industry. *Int J Appl Bus Econ Res* 15(11):311–334
21. Syal S (2015) Role of MSMEs in the growth of Indian economy foreign direct investment (FDI). *Pol De-reserv* 4(5):40–43
22. Tanriverdi H (2005) Information technology relatedness, knowledge management capability, and performance of multibusiness firms. *MIS Q* 311–334
23. Wang X, Xu M (2018) Examining the linkage among open innovation, customer knowledge management and radical innovation: the multiple mediating effects of organizational learning ability. *Baltic J Manag*
24. Wickert A, Herschel R (2001) Knowledge-management issues for smaller businesses. *J Knowl Manag* 5(4):329–337. <https://doi.org/10.1108/13673270110411751>

# Empirical Analysis of Static and Dynamic Stopword Generation Approaches



Saziya Begum Saiyed and Priti Sajja

**Abstract** In natural language processing, stopwords are considered as less meaningful words and contribute to reducing the speed of language processing. Stopwords are removed to ensure fast text processing and to decrease the size of the text. In this paper, experimental analysis of two different stopwords list generation approaches—static and dynamic are discussed. In the static approach, we have implemented the classic approach, and in the dynamic approach, we have implemented statistical Zipf’s law for automatic identification and removal of stopwords list. We have used a news dataset of 2225 documents for five different categories. We have compressed average document size by 44.53% using the static approach and 52.53% using the dynamic approach. This paper also shows the advantages and limitations of using both approaches.

**Keywords** Natural language processing · Stopwords · Classic approach for stopwords removal · Zipf’s Law · Automatic stopwords identification

## 1 Introduction

The world of digital documents is growing rapidly, and most of the digital libraries are in text format. To process these texts, one requires some natural language processing (NLP). NLP is a field of artificial intelligence that encloses a system for understanding and analyzes the text.

Text summarization is the branch of NLP that automatically creates a shorter version of the text (summary) which saves efforts and time of reading [1]. Extractive summarization is a type of summarization that decides important paragraphs, lines,

---

S. Saiyed (✉)

S.S.Agrawal Institute of Management and Technology, Navsari, Gujarat, India

P. Sajja

G.H.Patel PG Department of Computer Science and Technology, Sardar Patel University, Vallabh Vidhyanagar, Anand, Gujarat, India

e-mail: [priti@pritisajja.info](mailto:priti@pritisajja.info)



and words from original text and clubs them as a summary by using statistic and linguistic-based approaches [2].

Text preprocessing plays a vital role in making text usage ready for the summarization process. Preprocessing converts the original text into a predictable form, and that can be analyzed for further processing. The main tasks of preprocessing include tokenization, stemming, and stopwords removal.

Stopword identification and removal is also important in the task of text classification and sentiment analysis because it ends up with fewer terms than before, so it can improve the accuracy of classification/analysis.

1.1 About Stopwords

Stopwords are a crucial part of preprocessing. Stopwords are words in the text that does not have any significant value [3], and they occupy a large fraction of text in the document. They are found to consistently rule the corpus with no logical commitment. These words have no significant information, but they are used to satisfy grammatical syntax [4]. Words like articles, pronouns, adverbs, prepositions, conjunction, and injunctions are stopwords [5], and they are used for the construction of language.

Table 1 shows a sample general stopwords list. It can vary domain by domain. The ten most usually used words in English normally report 20–30% of the words in a document [6]. So, it is usually important to eliminate all stopwords terms for further Processing like analysis or classification of text.

There are two ways to find stopwords. The first way is to prepare a list of stopwords manually and match all words with that list and the second way is to prepare a list of stopwords automatically without using any corpus. A former approach is a static approach and the latter approach is the dynamic approach. A list of stopwords is needed in applications of data mining, information retrieval, automatic text summarization, text classification, NLP, etc.

In this paper, we have discussed static and dynamic stopwords generation. We have also performed an empirical analysis of static and dynamic approaches to find stopwords. Static approach uses a precompiled list of stopwords is prepared manually. First, the text is converted into words and then removes the word if it exists in the precompiled list of stopwords. Dynamic approach is aggravated using Zipf’s law [7] which affirms that for given linguistic, frequency of any word is inversely proportional to its rank in the frequency table. The dynamic approach does not need any

Table 1 Stopword examples

Stopword type	Stopwords
Determinants	a, an, the, another
Prepositions	Under, in, before, after, towards
Conjunctions	but, yet, so, or, for, nor, and

precompiled list, and it is used to generate automatic stopwords list and remove them from the text for further text processing.

We have applied Zipf's law for the dynamic generation of stopwords, which can be used in the corpus-free text.

## 2 Existing Work

Kaur and Sharma [2014] used Punjabi language documents of five categories as agriculture, entertainment, politics, sports, and health [8]. They manually developed a stopwords list for 1500 words. They concluded that an average 34.44% of stopwords were found in agriculture documents, while 32.70% in the entertainment category, 27% in health category, 30.45% in politics, and 33.62% stopwords in the sports category.

Stefano Ferilli et al. [2014] explained the methodology to learn linguistic resources automatically by applying high-level processing on documents [9]. To remove stopwords from a document, they prepared a list containing possible function words in English, French, Italian, Dialect, and Latin languages. They obtained an average 91.28% reduction of document for all languages.

Rakholia and Saini [2018] mentioned the term frequency-based statistical Zipf's law approach for information retrieval [10]. They applied Zipf's algorithm for 272 documents from the traveling corpus which were written in the Gujarati language. They concluded that only medium terms frequency words will be considered for further processing. They obtained a slope of log-log plot to linear fit as  $-0.997295926$  which is almost closer to  $-1$ .

Shah and Bhadka [2019] suggested the classic stopwords identification and removal approach and statistical Zipf's law approach [11]. They reduced file size average 23.09% for the Gujarati language using the classic approach.

## 3 Methodology

We have used the BBC news dataset and implemented static and dynamic approaches to remove stopwords from files of the dataset.

### 3.1 Dataset

We have used the BBC news dataset [12] of 2225 documents with five different categories. Business category contains 510 documents, entertainment category contains 386 documents, politics category contains 417 documents, sports category contains

511 documents, and technology category contains 401 text documents. Category-wise documents are merged to get five large documents to identify and remove stopwords from them.

### 3.2 *Static Stopword List Generation*

This approach is the classic approach which uses a precompiled list of Stopwords. We have used natural language tool kit (NLTK) stopwords list. A detailed algorithm is given below:

1. *Read the text from the document.*
2. *Obtain Precompiled list of Stopwords in the list named SWL.*
3. *Tokenize the text.*
4. *Apply stemming on tokens to find root words and store them in the list named L.*
5. *For each word in L*  
*Match word with SWL, If matched, then store the word in New List (NL) and remove it from the list (L).*
6. *Find the frequency of each unique word for NL.*
7. *Display each word from NL with its frequency.*

### 3.3 *Dynamic Stopword Generation*

This approach is inspired by Zipf's law. Zipf's law is a distribution model for the prediction of the most relevant term to the document. It is helpful when a very small number of the term occurs more frequently than a large number of terms occurs frequently in the text. It gives statistics about most occurring terms, infrequent terms, and relevant terms. To generate automatic stopword list and remove them, the algorithm is given below:

1. *Read the text.*
2. *Tokenize the text to get tokens.*
3. *Apply stemming on tokens to find root words and store them in the list named L.*
4. *Calculate the frequency of each word for L and store them in frequency table  $F(\text{word}, \text{frequency})$*
5. *Sort F based on frequency.*
6. *Remove the terms from head and tail(stopwords) from F.*

**Table 2** Results of static stopwords list generation approach

Category	Documents merged	Total terms	Unique terms	Stopwords identified	Non-stopwords identified	Compression rate (%)
Business	510	160,131	10,783	66,958	93,173	41.81
Entertainment	386	122,976	11,026	53,530	69,446	43.53
Politics	417	184,949	10,612	84,250	100,699	45.55
Sports	511	162,804	10,014	75,893	86,911	46.62
Technology	401	195,189	11,372	88,068	107,121	45.12

## 4 Experimental Results and Comparisons

### 4.1 Results

The result shows total number of terms, total number of unique terms, total number of stopwords, and total number of non-stopwords identified, and percentage of text compression. By using the static approach, 44.52% text compression is achieved and 52.53% text compression is achieved using the dynamic approach. Percentage of compression in dynamic stopwords list generation is dependent on the threshold of head and tail (i.e., percentage of terms removed from head and tail to generate the non-stopword list).

The results of the static stopwords list generation approach is shown in Table 2.

Table 2 shows the statistics of identified stopwords, non-stopwords, and compression rate after removing stopwords for five categorical documents for the static stopwords list generation approach.

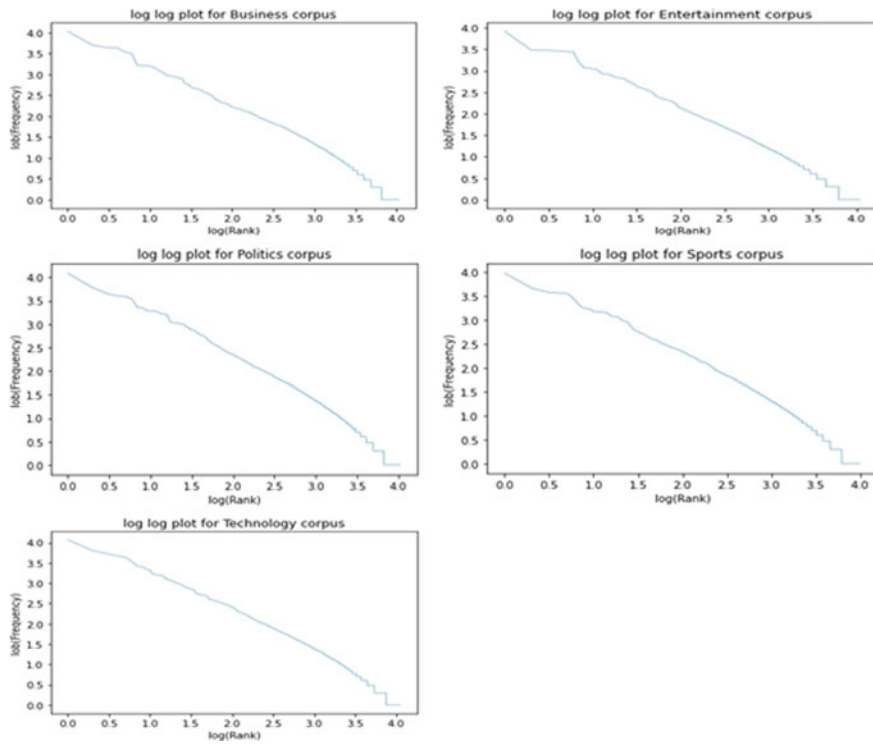
Figure 1 shows log-log plotting for our dataset using the dynamic stopwords list generation approach. As per Zipf's law graph of  $\text{Log}(\text{Frequency})$  against  $\text{Log}(\text{Rank})$  is roughly linear with the slope close to  $-1$  which excludes head and tail in plotting (as they are not linear).

Based on experiments, Table 3 shows that the slope is linear and nearer to  $-1$  after removing the approximate 37% terms from head and 25% terms from tail.

Table 3 shows statistics of several documents, number of unique terms and slope before and after removing head and tail terms. It also shows number of stopwords, number of non-stopwords and compression rate of a document after removing stopwords.

### 4.2 Comparison of Results

The comparison shows that the dynamic approach identifies more stopwords as compared to the static method in all five categories. But for the dynamic approach, it depends on the threshold decided with the help of the log-log plot shown in Fig. 1.



**Fig. 1** Log-Log plot for categorical documents

Table 4 shows the number of stopwords and the number of non-stopwords identified using the static and dynamic approach for five categorical documents.

## 5 Conclusion

The dynamic approach identifies and removes more stopwords than the static approach. It is worth noting that this number also depends on the threshold decided for head and tail in the case of the dynamic approach. The static approach compares each word of text with a precompiled list of stopwords, so it takes time for comparison; hence, it has the disadvantage of an increase in processing time when text size increases. The dynamic approach does not use the precompiled list of stopwords; stopwords are removed automatically without making any comparisons by implementing Zipf's law. Saving of comparison time is an advantage in the dynamic approach. The limitation dynamic approach is occasionally non-stopword is considered as stopword and also removed when we remove the most frequent and rare words from the list.

**Table 3** Results of dynamic stopword list generation approach

Category	Documents merged	Total terms	Unique terms	Slope before removing head and tail	Slope after removing head and tail	Stopwords identified*	Non-Stopwords identified <sup>a</sup>	Compression rate (%)
Business	510	160,131	10,783	-0.75	-1.04	82,630	77,501	51.6
Entertainment	386	122,976	11,026	-1.17	-0.97	67,618	55,358	54.98
Politics	417	184,949	10,612	-1.31	-1.00	96,921	88,028	52.4
Sports	511	162,804	10,014	-1.30	-1.02	81,227	81,577	49.89
Technology	401	195,189	11,372	-1.30	-1.00	104,982	90,207	53.78

<sup>a</sup> After removing terms from head and tail

**Table 4** Comparisons of results

Approach			Static stopword list generation		Dynamic stopword list generation	
Category	Total terms	Unique terms	Stopwords identified	Non-stopwords identified	Stopwords identified	Non-stopwords identified
Business	160,131	10,783	66,958	93,173	82,630	77,501
Entertainment	122,976	11,026	53,530	69,446	67,618	55,358
Politics	184,949	10,612	84,250	100,699	96,921	88,028
Sports	162,804	10,014	75,893	86,911	81,227	81,577
Technology	195,189	11,372	88,068	107,121	104,982	90,207

## References

1. Saziabegum S, Sajja PS (2016) Literature review on extractive text summarization approaches. *Int J Comput Appl* 156(12)
2. Gupta V, Lehal GS (2010) A survey of text summarization extractive techniques. *J Emerg Technol Web Intell* 2(3):258–268
3. Tijani O, Onashoga S, Akinwale AT (2017) An auto-generated approach of stop words using aggregated analysis
4. Zou F, Wang FL, Deng X, Han S (2006) Evaluation of stop word lists in Chinese language. In: *LREC*, pp 2497–2500
5. Cook Cooke N, Gillam L (2008) Distributional lexical semantics for stop lists. In: *BCS-IRSG workshop on corpus profiling*, pp 1–11
6. Chekima K, Alfred R (2016) Automatic construction of malay stop words based on aggregation method. In: *International conference on soft computing in data science*. Springer, Singapore, pp 180–189
7. Zipf GK (1949) *Human behavior and the principle of least effort*. Addison-Wesley, Cambridge, (Mass.), p 573
8. Kaur R, Sharma S (2014) Pre-processing of domain ontology graph generation system in Punjabi. *arXiv preprint arXiv:1411.5796*
9. Ferilli S, Esposito F, Grieco D (2014) Automatic learning of linguistic resources for stopword removal and stemming from the text. *Procedia Comput Sci* 38:116–123
10. Rajnish R, Jatinderkumar S (2018) Impact of Zipf's law in information retrieval for Gujarati language. *Int Arch App Sci Technol* 9(2):36–40
11. Shah DN, Bhadka H (2020) Noise removal as pre processing task and its implementation for gujarati named entity recognition. In: *ICT for competitive strategies: proceedings of 4th international conference on information and communication technology for competitive strategies (ICTCS 2019)*, December 13th–14th, 2019, Udaipur, India. CRC Press, p 275
12. Pariza Sharif. *BBC News Summary, Version 2* (2018). Retrieved on 19/1/2021 from <https://www.kaggle.com/pariza/bbc-news-summary>

# Rethinking Access Control Mechanism for Ubiquitous Computing



Nalini Mhetre , A. V. Deshpande, and Parikshit N. Mahalle 

**Abstract** The Ubiquitous computing (UbiComp) offers an ever-present computing environment. It is characterized by ever-growing small, networked, and portable everyday objects having computational capability along with conventional computers. Dynamic usage scenarios of several devices generate new associations. These devices may not have past information of each other and a common trusted third party may be absent. Traditional access control mechanisms rely on a previous knowledge of devices to grant or deny access to resources making them unsuitable in a UbiComp environment where entities are not foreknown to each other. This set up new challenges for resource security. In this paper, a decentralized access control mechanism is proposed for UbiComp to secure its resources. The proposed access control mechanism incorporates the high-level context and the trust level of the device requesting a resource. We also suggest the use of a machine learning approach to derive high-level context. Proposed mechanism considers device type as a high-level context derived through clustering.

**Keywords** Access control · Authorization · Context-aware access control · Trust model · Ubiquitous computing

## 1 Introduction

UbiComp applications are everywhere and probably we don't even notice them [1]. But that's how they are supposed to work without intruding, e.g., Apple Watch, Smart speakers like Google Assistant, etc., there is a very wide variety of devices that are compatible with UbiComp systems ranging from traditional computing

---

N. Mhetre (✉)

Department of Computer Engineering, Sinhgad College of Engineering, Pune, India

A. V. Deshpande · P. N. Mahalle

Department of Computer Engineering, S.K.N. College of Engineering, Pune, India

e-mail: [principal.skncoe@sinhgad.edu](mailto:principal.skncoe@sinhgad.edu)

P. N. Mahalle

e-mail: [aalborg.pnm@gmail.com](mailto:aalborg.pnm@gmail.com)



devices, Smart phones, smart objects to sensors. Most of these devices have different computing powers, input interfaces, operating systems, visual displays and run on different communication protocols. Some are designed for end-user interaction, whereas other devices, such as sensors, keeps running without any end-user interference; hence, heterogeneity is the primary characteristic of UbiComp. Applications of UbiComp have scenarios where frequent change is there in devices, networks, and software components. Spontaneous interaction between devices often implies that these devices have no or little prior information of each other and a common trusted third party may be absent. Due to such a volatile nature of the system, trust is always lowered and the security of devices and their resources are the main concerns.

As objects interact across multiple organizational boundaries, adaptive services and applications have emerged. Hence, there is a requirement for context-aware adaptive security. Further, it is also important to identify whether the communicating object is an adversary or not, i.e., level of trust. UbiComp is more characterized by security versus functionality issues. Interconnected objects share various resources and critical information with each other and are bound to some access control mechanisms. Guaranty of users privacy, confidentiality, integrity and accessibility are basic security requirements for granting access to resources. Access control is a basic security method that regulates the use of resources in systems boundaries to protect them against unauthorized access.

Access control models like access control lists (ACL), rule-based access control, role-based access control (RBAC) [2] are designed for traditional computing environments, considering systems have prior knowledge of objects and their identities. However, these access control models are static in nature and they make use of objects' static attributes to do access control decisions. They are not capable to acquire contextual information from the objects and their surrounding environments. Consequently, they are unable to respond to changes in the contextual attributes which might have security implication. Static access control models such as these are not appropriate for UbiComp's dynamic and ad-hoc nature. Therefore, an adaptive context-aware access control mechanism with trust management is an immediate need.

So far, some of the attempts have been made to model context-aware access control. In addition to traditional systems, context-aware systems make use of contextual information like locations, time, behavior, environmental conditions, network traffic data, device attributes, etc. Recently, to address the reliability of contextual information, users, context consumers, context providers, and services, trust management models are demonstrating themselves as powerful tools [3]. In addition, users should be able to gain more control over their device's capabilities by restricting those application permissions in sensitive context.

We propose here a context-aware access control model using device trust to secure resources in the UbiComp environment. This model utilizes device type as a high-level context as UbiComp has the potential to use context awareness mechanism to build smarter and dynamic access control systems using operational and huge data of variety of smart devices.

Other sections of the paper are organized as follows: motivation is discussed in the Sect. 2. Section 3 describes related work. The details of proposed approach are described in Sect. 4. And, Sect. 5 concludes the paper.

## 2 Motivation

Traditional access control systems consider device preferences. They assume that these preferences are unchangeable as things or objects move from current location to a different location. In addition to traditional systems, context-aware systems use contextual information like locations, time, behavior, environmental conditions, network traffic data, device attributes, etc. Even though these models have widespread success, but still their scope is limited in the contextual information type they use and inferred knowledge from the context information. Managing so many devices along with their huge data is really challenging. Trust management is also important to address the reliability of devices and their surroundings.

Smart Cities and Communities have gained lot of attention in last couple of years. Concepts like convergent access control [4] are surfaced on top of traditional access control frameworks which are focused on large and diverse systems. Dynamic authorization, flexibility, scalability, decentralization, Light weight and privacy preserving are essential requirements for new systems. Thus, it is significant to rethink current access control mechanism for such heterogeneous, diverse and dynamic ecosystem.

The security solutions of traditional and generic computing devices are not adequate in a smart UbiComp environment. The UbiComp environment has too many challenges like management of a huge number of devices, many of these devices are not capable of running security solutions, handling massive amounts of data, etc. that is why the use of machine learning methods can help. Machine learning complements human efforts and sometimes can even replace humans. It skillfully finds and creates patterns to improve UbiComp security, especially when huge amounts of data are fed. And already there is a huge amount of data available in the UbiComp environment.

## 3 Related Work

### 3.1 Background

ACL is most common form of access control, which is inefficient and complexity arises when the count of subjects and resources increases. To increase efficiency of basic ACLs, RBAC model was designed. RBAC created an abstraction layer between subject and resource and took the burden of assigning one to one access rights by creating roles which can be assigned to multiple subjects. RBAC model is

a centralized architecture where a central entity is responsible to manage all access control requests based on their authorization policies. Due to increased complexity in systems and the number of roles, RBAC becomes inefficient for highly distributed network environments. To overcome the drawback of RBAC model, attribute-based access control [ABAC] [5, 6] model and organizational-based access control [OBAC] [7] model were introduced. ABAC and its variants are more manageable and scalable compared to role-based models. Difficulty in enforcing least privilege access principle is a common problem of ACLs, RBAC, ABAC and OBAC systems.

Context-awareness has emerged in discussions in various fields these days. However, most of today's context-sensitive systems and applications are still in their early stages and use only atomic context information, like the location, time and identity. According to Dey, *"a system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the users' task"* [8]. Context information can be categorized in many possible ways but operational and conceptual categorization are most widely used. Henricksen and Indulska classified context into 4 categories based on how context is captured [9]: sensed, static, profiled, or derived.

Table 1 shows summary of related work indicating the parameters: extensibility of model, use of context and trust along with its core concepts and issues. Extensibility of the access control model takes into account constraints dynamically generated due to addition of new devices considering underlying smart environment. In Table 1, value Y means that model considers this parameter and N means model does not consider this parameter.

**Table 1** Summary of related work

	Extensible	Context-aware	Trust	Concepts
[3]	Y	N	Y	<ul style="list-style-type: none"> <li>Trust level is the key concept</li> <li>Distributed model for IoT</li> </ul>
[6]	Y	N	N	<ul style="list-style-type: none"> <li>Enhanced model with combination of RBAC and ABAC</li> <li>Flexible and granular model</li> <li>Suitable for both cloud and IoT</li> </ul>
[10]	Y	Y	N	<ul style="list-style-type: none"> <li>Centered on the notion of capability</li> <li>Hybrid model for IoT Domain</li> </ul>
[11]	Y	N	Y	<ul style="list-style-type: none"> <li>Trust aware access control model based on fuzzy logic</li> <li>Scalable and energy efficient</li> <li>Distributed access control model</li> </ul>
[12]	N	Y	N	<ul style="list-style-type: none"> <li>The notion of context to match the IoT requirements</li> <li>Less complex compared to traditional access control models</li> </ul>
[13]	Y	N	N	<ul style="list-style-type: none"> <li>Smart contacts are used for storing access control matrix</li> <li>Extended RBAC model</li> <li>Computational overhead</li> </ul>

### 3.2 *Gap Analysis*

Table 1 shows access control models for IoT and UbiComp. It shows need of a more precise model that takes into account some issues such as extensibility, context-aware and trust. As UbiComp is distributed and decentralized in nature, there is a need to make access control mechanisms proactive especially for derived context which is missing. In this view, there is a need to associate machine learning techniques (supervised/unsupervised) according to underlined application with access control mechanisms. The summary shows that little efforts have been made in this area to consider the recommendation, experience and knowledge perceived through machine learning. Also, access control should be modeled on the trust relationship between resource requester and provider. Thus, there is a need for an access control mechanism which responds proactively to change in trust and context conditions over the time for the dynamic authorization adjustment.

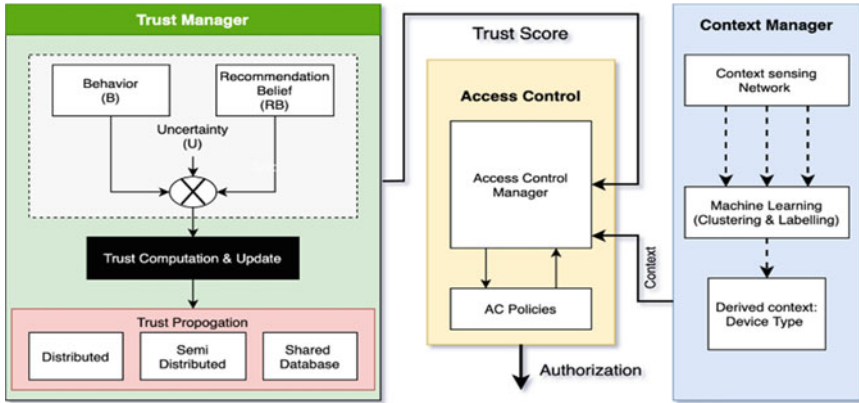
## 4 *Proposed Work*

To model access control architecture there are two Approaches: centralized architecture and distributed architecture. A central entity in centralized architecture is responsible to manage all access control requests based on their authorization policies. In some cases, access control decisions are influenced by partial involvement of end-devices. Whereas, in the distributed architectures, the smart end-devices have the access control logic. They are capable of taking authorizations decisions without any centralized entity. UbiComp is characterized by its dynamic nature in terms of mobility, heterogeneity and ad-hoc networking capabilities. Therefore, centralized architecture is not suitable for UbiComp.

In this section, we propose a trust and context-aware access control framework as depicted in Fig. 1. It is based on decentralized approach in which devices without a centralized authority may make authorization decisions. It has 3 main functional components, namely: trust manager, context manager, and access control module. These are explained further in this section.

### 4.1 *Trust Manager*

For the secure device to device communication in UbiComp, the communicating entities must be trusted. When device 1 (subject) requests some resource (like File, service, data, etc.) from device 2, it would be of great help for device 2 to know the trustworthiness of the subject for security reasons. It is of utmost importance while granting sensitive resources. In this framework, the trust manager component computes the trust of the subject. This framework proposes to use our previous work



**Fig. 1** Proposed trust and context-aware access control framework

suggested in [14] for the fuzzy trust score calculation of the subject. It considers the communication history (B: Behavior) of the subject with the object and peers recommended the trust value (RB: Recommendation Belief) of the subject to find its trust score. In the absence of sufficient history, there is often uncertainty about whether or not an interaction would succeed, which could also affect trust decision making. This trust value affects the access rights of the subject. It is one of the major points to be considered while designing access control policies. The relationship of trust in term of B, RB, and  $U$  is shown in Eq. 1.

$$T = f(B, RB, U) \quad (1)$$

Trust fuzzy values are labeled as low, average and high. To achieve high-level of precision, granularity of labels may be increased. Trust will be used for establishing end-to-end secure communication.

## 4.2 Context Manager

This module is responsible for investigating what contextual factors would affect while granting access control requests for the resources. The context sensing part in this module gathers a variety of information about the device and its environment that may play the role of contextual parameters according to the application or service in use. These parameters together represent the state of the device of a particular type. Thus, knowledge or learning of these parameters would be useful in identifying and detecting patterns proactively which represent the state of the device putting it in a particular logical category. The proposed framework suggests the use of machine learning to process context sensed to infer knowledge about the device

as suggested in our previous work [15]. This work uses network features or event information of the devices as a dataset for unsupervised machine learning, in our case k-means clustering. This model assists in identifying general trends of network parameters in devices in order to aggregate them under a specific class name. This inferred knowledge or cluster labels may be used as a high-level context that is a derived context indicating the type of device as **CONSTRAINED**, **POWERFUL**, or **SEMIPOWERFUL**, that requests for resource. Proposed access control mechanism makes use of this high-level or abstract context in designing access policies. Further, the number of clusters may be increased in future. The main advantage of the context manager module is that it scales well with ever-growing and unknown devices in the UbiComp. Thus, the value associated with the high-level context of the subject affect's authorization decisions.

### 4.3 Access Control Module

Access control module contains: 1. an access control manager, and 2. a set of access control policies. A policy consists of a collection of access rules. The Access Control module is responsible for: (a) Mapping of access rights with access rules, and (b) Authorization decision making. The access right and access policy considers Trust Score and Context for access permission. The proposed access control model relies on trust score and contextual data of the requesting entity to assign the access permission to the resources in the current situation. Proposed framework models access control (AC) as a function of trust ( $T$ ) and context ( $C$ ).

$$AC = f(T, C) \quad (2)$$

For example, semantic variables of access rights (AR) are listed as in Eq. (2) and the corresponding semantic variables of access permissions (AP) are listed in Eq. (3). It means that *Default* AR maps to *Basic* AP, *Expert* AR maps to *Moderate* AP, and so on. The total count of semantic variables of AR is equal to the count of AP.

$$AR = \{\text{Default; Expert; Advanced}\} \quad (3)$$

$$AP = \{\text{Basic; Moderate; Full}\} \quad (4)$$

This module can be easily extended to handle different situations by extending the AR variables. The access control module maps the AR to the AP; therefore, the number of AR variables must be equal to number of AP variables. This makes the model more scalable and dynamic to the change of context over time.

## 5 Conclusion

In this paper, the unsuitability of traditional access control models in UbiComp is discussed and the conceptual access control framework is proposed. The paper further elaborated on the components of the proposed framework. With development in new technologies, new types of devices emerge in the market and, hence, authorization and access control in UbiComp is device centric. Our proposed approach considers the use of trust relationships amongst devices and their high-level context in policy designing and decision making. Devices are enabled to obtain high-level context through machine learning by processing a rich dataset about the devices and their networking environment. The use of machine learning is a salient feature of the proposed framework and is quite suitable for UbiComp. The future scope of this work includes developing and analyzing a prototype based on the proposed framework.

## References

1. Weiser M (1999) The computer for the 21st century. *ACM SIGMOBILE Mob Comput Commun Rev* 3:3–11. <https://doi.org/10.1145/329124.329126>
2. Sandhu RS (1997) Role-based access control
3. Wang J, Wang H, Zhang H, Cao N (2017) Trust and attribute-based dynamic access control model for Internet of Things. In: *Proceedings of international conference cyber-enabled distributed computing and knowledge discovery 2017 (CyberC 2017)*, 2018-Janua, pp 342–345. <https://doi.org/10.1109/CyberC.2017.47>
4. Bhatt S, Sandhu R (2020) Convergent access control to enable secure smart communities. In: *Proceedings of the second IEEE international conference on trust, privacy and security in intelligent systems, and applications, TPS-ISA 2020*, pp 148–156. <https://doi.org/10.1109/TPS-ISA50397.2020.00029>
5. Kaiwen S, Lihua Y (2014) Attribute-role-based hybrid access control in the Internet of Things. In: *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 8710 LNCS, pp 333–343. [https://doi.org/10.1007/978-3-319-11119-3\\_31](https://doi.org/10.1007/978-3-319-11119-3_31)
6. Attia HB, Kahloul L, Benharzallah S (2018) A new hybrid access control model for security policies in multimodal applications environments. *J Univers Comput Sci* 24:392–416
7. Kalam AAE, Baida RE, Balbiani P, Benferhat S, Cuppens F, Deswarte Y, Mieke A, Saurel C, Trouessin G (2003) Organization based access control. *Proceedings of POLICY 2003 IEEE 4th international workshop on policies for distributed systems and networks*, pp 120–131. <https://doi.org/10.1109/POLICY.2003.1206966>
8. Abowd GD, Dey AK, Brown PJ, Davies N, Smith M, Steggles P () Towards a better understanding of context and context-awareness. In: *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 1707, pp 304–307. [https://doi.org/10.1007/3-540-48157-5\\_29](https://doi.org/10.1007/3-540-48157-5_29)
9. Henriksen K, Indulska J (2004) Modelling and using imperfect context information. In: *Proceedings of second IEEE annual conference on pervasive computing and communications workshops PerCom*, pp 33–37. <https://doi.org/10.1109/PERCOMW.2004.1276901>
10. Anggorojati B, Mahalle PN, Prasad NR, Prasad R (2012) Capability-based access control delegation model on the federated IoT network. In: *International symposium on wireless personal multimedia communications, WPMC*, pp 604–608

11. Thirukkumaran R, Muthukannan P (2019) TAACS-FL: trust aware access control system using fuzzy logic for Internet of Things. *Int J Internet Technol Secur Trans* 9:201. <https://doi.org/10.1504/IJITST.2019.098140>
12. Bouij-Pasquier I, Ait Ouahman A, Abou El Kalam A, Ouabiba De Montfort M (2016) SmartOrBAC security and privacy in the Internet of Things. In: *Proceedings ACS/IEEE international conference on computer systems and applications, AICCSA*, July 2016. <https://doi.org/10.1109/AICCSA.2015.7507098>
13. Xu R, Chen Y, Blasch E, Chen G (2018) BlendCAC: a smart contract enabled decentralized capability-based access control mechanism for the IoT. *Computers* 7:1–27. <https://doi.org/10.3390/computers7030039>
14. Mhetre NA, Deshpande AV, Mahalle PN (2016) Trust management model based on fuzzy approach for ubiquitous computing. *Int J Ambient Comput Intell* 7:33–46. <https://doi.org/10.4018/IJACI.2016070102>
15. Mhetre NA, Deshpande AV, Mahalle PN (2021) Device classification-based context management for ubiquitous computing using machine learning. *Preprints*. 202101027. <https://doi.org/10.20944/preprints202101.0273.v1>



# Review of Smart Building Management System



Mona Tatari, Parul Agarwal , M. Afshar Alam , and Jawed Ahmed

**Abstract** Smart buildings, use information and communication technologies for automating the process of heating, ventilation, cooling, security, water and energy management, health of elderly, and thus aids in achieving energy efficiency. There is an urgent need for smart energy-saving buildings, and modifying the existing buildings so that they are able to use various automation systems and smart energy management systems in an effective manner. This paper provides extensive and systematic review of building management systems (BMS) which can be used to improve building energy consumption, achieve energy efficiency, and reduce maintenance and repair costs. In this paper, some studies would be analyzed to identify the technical advances that have laid the base for a better future. This paper also presents the challenges that confront the real-life scenarios and proposes a guideline for the proposed solutions. It can be concluded that various machine to machine and advanced technologies can enable a building to respond intelligently to achieve energy efficiency.

**Keywords** Building management system (BMS) · Information and Communication Technologies (ICT) · Smart Building · Smart energy · Energy consumption · Smart technology · Smart city · Energy efficiency · Sustainable · Energy management · Artificial intelligence · Green energy

## 1 Introduction

All buildings, whether commercial or residential, old or modern, public or private, are places where people live and work and have fun. Buildings are the largest consumers of energy. Where energy consumption in buildings represents more than 40% (in India), this percentage may rise to almost 60%. The buildings use various energy services on a daily basis to ensure the comfort and security of their residents. Hence, buildings are responsible for three quarters of all urban greenhouse gas emissions.

---

M. Tatari · P. Agarwal (✉) · M. A. Alam · J. Ahmed  
Department of Computer Science and Engineering, Jamia Hamdard, New Delhi, Delhi, India  
e-mail: [pagarwal@jamiahamdard.ac.in](mailto:pagarwal@jamiahamdard.ac.in)

Hence, the need to make buildings smart, green and sustainable energy [1]. In view of the environmental, economic and political reasons, there is a need to reduce the energy consumed in buildings. Among these reasons that require reducing energy consumed in buildings are the high percentage of carbon dioxide in the atmosphere, the need to get rid of fossil fuels and stabilize energy networks in buildings [2]. The smart city concept integrates several areas including smart government, smart transportation, smart economy, smart people, smart environment and smart life. The buildings sector has the most role in adapting to new challenges in all previous areas. A sustainable smart city is considered a modernized city that uses information and communication technologies (ICT) and other means to improve life services and operational efficiency, so as to ensure that all current and future economic, social, environmental and cultural needs are met without affecting the lives of residents [3]. A smart building is a place that uses modern information and communication technology for automation systems to ensure the safety and comfort of residents, efficient energy consumption and efficient job scheduling for construction activities. A building management system is a control system that monitors and controls the mechanical and electrical building services [4]. The primary goal of energy management in buildings is to reduce used energy and reduce electricity bills and increasing passenger comfort without affecting living standards [5]. Section 2 discusses the related papers, the techniques they used and summarizes a few of them with respect to the component of smart city under review by those papers. Smart building management systems faces a few obstacles associated with its adoption and implementation in real-life. Section 3 discusses the challenges and solution which can be adopted to lead a sustainable life. Section 4 presents the conclusion of the paper and proposes the future work that can be carried out.

## 2 Literature Review

Smart building is defined as a set of communication and information technologies that help connect, manage, control and automate a variety of different systems that work independently within the building from a distance. In this section, we present the previous published literature in the field of smart buildings, their technologies, and the applied energy efficiency methods. In [2], an intelligent building energy management system (iBEMS) that uses open communication protocols has been proposed in shopping malls. The proposed system consists of a central system that communicates with all subsystems within the building, and exchanges information in real time using open communication protocols. iBEMS uses input from installed sensors and external data and makes appropriate operating decisions. It also provides drawings showing the entire building with all its subsystems. In [3], the methodology for integrating smart buildings in the smart city ICT platform (SBISC) was presented. The analysis showed that the environment and smart energy are the area's most conducive to integrating smart buildings into the smart city platform. The proposed system helps interoperability between the smart building and the external digital

platform, and focuses on the use of artificial intelligence techniques in managing the smart building and the smart city. In [4], they implemented a smart energy management system (SBMS) by using a private cloud implemented with the open source Eucalyptus 2.0.3 binary packages for the CentOS OS in the Computer Systems Engineering Laboratory at the University of Oradea. Eucalyptus is a flexible and easy to install program that uses standard communication protocols to coordinate the work of all different web services with each other. In [5], they discussed building and home automation networks, wireless sensors and motors based on Zigbee technology. They discussed the hardware power management diagrams. They discussed smart lighting control schemes based on wireless actuator and sensor networks. In [6], an intelligent building energy management system using renewable energy sources and Zigbee technology is proposed. Renewable energy sources are connected to building using a battery and an inverter, renewable energy sources is stored for use during peak hours. In this method, the traditional fixed pricing models are canceled, and new models are used that depend on the pricing of the time of use. In [7], a smart grid test architecture has been proposed for use on AUI campuses. The proposed smart grid consists of power systems, control systems, and an information and communication technology component. The ICT component consists of: the wireless sensor network, the wireless operator networks, and Big Data Analytics Platform (BDAP). The wireless sensor network captures data from different kind of sensors installed. This data is processed in the data analysis platform, the required control decision is made, and BDAP issues orders to the operator network in order to further enhance energy efficiency via context awareness. In [8], they implemented a prototype involving a single user with controllers in two buildings (home building and office building). They collected energy consumption data for electrical appliances, for a period of one year, then evaluated and analyzed them to determine the pattern of energy consumption. They proposed to use Internet of Things technology to smartly control the energy use of devices. In [9], a building management systems (BMS) with IOT and wireless communications technology in Nigeria, which the need for more energy efficient smart buildings. Because of the Energy efficiency and economic benefit of BMS, hence the need to promote the use of this technology in a developing nation such as Nigeria. In [10], a smart building automation control unit was designed and implemented based on the use of artificial intelligence technology (AIBSBAC) and was used to operate commercial and residential building automation systems. The proposed design consists of subsystems that define the smart user, subsystems that improve the internal and external environment, and a subsystem for decision-making using artificial intelligence and a global infrared communication system. In [11], it is proposed a building management system based on artificial intelligence tools for more advanced data processing, the sophisticated processing of sensor-generated data which makes the smart building into a “Cognitive Building.” An AI platform helps for the introduction of intelligent learning processes. This can be either cloud-based or server-based. Cloud-based server farms offer more processing power, and cloud-based AI frameworks offer a broader range of features. The AI platform is built on a smart building infrastructure, and all technical systems should ideally be connected to that must be able to govern the building facility and room automation systems. In [12],

the ML algorithms used in intelligent monitoring are reviewed and they proposed the Interpretable Artificial Intelligence (XAI) approach, which helps bypass the hurdles of not understanding ML algorithms. Artificial intelligence algorithms, combined with traditional technologies, help process large amounts of data. To increase confidence in the use of AI smart surveillance systems, the “black box” nature of AI algorithms for intelligent surveillance was proposed, which led to the concept of “interpretable artificial intelligence.” In [13], they describe the differences between the applications of modern smart technologies and traditional technologies used in commercial buildings, and the amount of energy saved by using smart technologies. Information and communication technologies used in smart buildings contribute to the control and automation of all systems in them. It also helps to increase the safety and comfort of the occupants by using less energy compared to traditional buildings. In [14], they indicated that energy efficiency come from Smart lighting and Smart HVAC. The lighting levels are adjusted according to the times, considering the air conditioning, heating and ventilation systems in the building and the design of the windows, in addition to the use of occupancy sensors. Dali or Modbus, is limited to the lighting control protocol. Smart HVAC communicate with different sensors that work according to weather forecasts, occupancy rate, and the ineffective system to quickly and automatically adjust and control these systems. For heating and cooling applications Opentherm is designed, and MBus for building metering. In [15], they proposed an intelligent model for energy-saving buildings using Internet of Things technology to manage all technical systems in the building. They also proposed a remote-control method with the use of a cloud interface to improve the energy performance of buildings. IoE technology, which is an application IoT technology, has been proposed to be used in the management of smart energy systems, achieving energy efficiency. In [16], they talked about the use of IoT technology in most smart buildings in order to link the work of building systems together using different sensors or IoT devices connected to the network. Using IoT in smart buildings, all connected building systems can be controlled and monitored remotely from anywhere. In Table 1, we review the smart city components and the techniques used which have been discussed in the previous papers.

The following present the papers which suggested models or presented algorithms and contributed in making their buildings more intelligent and achieving energy efficiency. In [17], the authors proposed a smart system consisting of various sensors and actuators linked to a central system to coordinate their work, a power box that acts as a local building management system, and temporarily stores data needed, the data is sent to a statistical server and stored to make the necessary statistics, and then, the appropriate data is extracted. They also proposed, EA-SIM [18] which is a new, open source, and continuous framework for designing and simulating energy-aware smart spaces that rely on System C TLM-2.0 as a representative abstraction to test policies and predict environmental behavior. In other work [19], the authors were created a model consists of 5 smart buildings. Each smart building has loads, energy storage, a small renewable source (PV) and a building automation. A demand response program was applied for energy management that depends on the response

**Table 1** Smart city components and the techniques used

References	Components of smart city	Technique
[2]	Shopping malls	Intelligent building energy management system (iBEMS) with supported communication protocols: Modbus, BACnet, LonWorks, ZigBee, KNX, DALI, SMI, EnOcean
[3]	Buildings	Smart building integration into the smart city ICT platform
[4]	OradeaUniversity campus	SBMS based on a private Cloud Implemented with Eucalyptus Framework
[5]	Buildings: Appliances and lighting	Building automation networks based on Zigbee technology Non-intrusive appliance load monitoring (NALM) technology Intelligent WSN-based lighting control schemes
[6]	Home building	Smart building energy management system using renewable energy and Zigbee technology
[7]	AUI university campus	Smart micro- grid with information and communication technology (ICT): IOT and Big Data Approach
[8]	The user's office building and home building	IoT networking and computing technologies
[9]	Commercial buildings	Building management systems (BMS) with IoT and wireless Communications technology
[10]	Commercial and residential buildings	Artificial intelligent-based smart building automation controller (AIBSBAC) based on a fuzzy logic
[11]	Buildings	Linking building automation with a (cloud-based) AI platform
[12]	Smart infrastructure	Explainable artificial intelligence (XAI) technical discipline
[13]	Commercial buildings	HVAC, Smart thermostat, Wired sensor, Wireless sensor. Plug load: Smart plug, Advanced power strip. lighting: Advanced lighting controls, Web-based lighting mangement system, Window Shading:Automated shade system, Switchable film, Smart glass. Building Automation: BAS Analytics: Cloud-based energy, information system (EIS). DER: Smart inverter

(continued)

**Table 1** (continued)

References	Components of smart city	Technique
[14]	Buildings	Common standards: Wi-Fi, Bluetooth, Zigbee, EnOcean Thread or KNX. Lighting control Protocol Modbus or DALI gate is used, heating and cooling applications is used Opentherm And for metering MBus
[15]	Buildings	IoE Internet of Things technology which includes utilizing smart sensors and renewable energy integration
[16]	Buildings	Internet of Things technology (IoT). Networked IoT sensors or devices

of the price signal value and directly controls the loads from the facility. The simulation was performed using any logic 6 [20], and every smart building is simulated as an agent. Several scenarios were used to analyze. In another study [21], they presented an integrated improvement model that simulates an intelligent energy management system that includes decisions related to the heating and cooling system in addition to decisions about available energy sources. All processes in conventional and smart BEMS systems are designed using an optimization model [22]. The integrated model considers the energy balance of a strategic model [23, 24] and a full treatment of building physics by lower operating units [22–25]. They took data from two buildings which are located in the temperate climate zone of Spain and Austria. In another paper [26], the authors proposed a model based on modern information and communication technologies such as clouds of things, Internet of Things, and advanced artificial intelligence. The information expected from user behavior is collected from smart home systems. The AAI advanced artificial intelligence system controls and processes all smart home services. IoT and CoT technologies are used to collect all AAI systems. An infrastructure is created based on the information and communication technology of the smart city. Another paper [27], presented the use of IoT platforms in a smart building. The authors have proposed a thermodynamic model that uses the human body heat balance equation with the thermodynamic properties of the users. The proposed model was implemented in two smart buildings and used heaters that could be controlled with a temperature indication and thermal comfort using MATLAB/Simulink®. In [28], the authors have designed and simulated a topology system helps save energy in a smart building, uses the sensors that collect and control data, and a cloud server that manages the building and stores only valid information coming from the sensors. The existing network uses communication protocol IPv6 and transmission protocol (NAMRTP). The simulation is suggested using Contiki operating system (OS) [29] which is an open source system used to collect a large amount of data. Then, it is stored in special files for future analysis.

### 3 Challenges and Solutions

Smart buildings are characterized by many advantages that we mentioned previously, but they face some challenges and obstacles that hinder their adoption and implementation so in this section we present the Challenges and Solutions associated with it. In [2], they proposed an intelligent system consists of a central system that communicates with subsystems within the building and exchanges data in real time and uses open protocols in building communication. Consumed energy can reduce by using this system compared to a traditional BMS system that use multiple communication protocols. In another paper [3], the challenge was how to align the smart building concept with strategic improvement plans in a smart city, so they proposed an Smart Building in a Smart City (SBISC) that reveals the potential for interoperability between the smart building and smart city, with a focus on future directions of artificial intelligence in managing both the smart building and the smart city. In [5], they discussed privacy issues, as the use of smart technologies in buildings causes a risk to the security and privacy of its residents, which hinders the deployment of smart meters. So, they suggested a digital rights management system (DRMS). The utility is allowed monthly access to customer usage data, but they are required to grant permission to reach their data with greater levels of accuracy in exchange for discounts or other incentives. In [9], they discussed several factors which have resulted in BMS limited applications in Nigeria that include: BMS is still not a widely accepted technology. The level of awareness as regards the use and benefits of BMS is quite low. Electricity is inadequate. Most building equipment suffers from maintenance. The state of Nigerian economy is poor. By evaluating the economic advantages and energy efficiency benefit of utilizing BMS in commercial buildings in Nigeria, a lot of awareness will be created as to the need for Nigerian private and public developers to embrace the use of this technology. In [15], they proposed a smart building model that helps save energy by using IoT technology to manage all the technical systems in the building. There were several limitations including the equipment and technical tools that should be used, financial limitations in order to apply the simulation to the proposed model of smart building and its management system. They proposed to implement a management system that controls and tracks the proposed smart building model, and to test a management system that achieves energy efficiency for the current building. In [26], they discussed the challenges related to sustainability solutions in need due to the great shortage of resources, the high population density, which causes an increase in the consumption of existing resources. They proposed a more advanced artificial intelligence (AAI) system that links all artificial intelligence systems in smart homes using internet and cloud of things technologies. In [30], the authors analyzed the main economic, political, social, and technical barriers to smart building adoption in Brazil based on the literature review and expert interviews. They suggested some solutions include: Preparing technical courses. Developing cost-sharing plans for infrastructure installation. Using low-cost energy generation technologies. Adopting

smart building technologies in social housing programs. An increase and development in the infrastructure investments of the electrical system. Providing low-income energy users with individual smart meters at competitive prices. Thus, we see that though several challenges are associated with smart building management, yet, effective solutions and technological advances [31] can be utilized in smart buildings for a sustainable future.

## 4 Conclusion

In this paper, we conducted an extensive literature review to identify all technologies approaches that could be applied in smart buildings. The development of communication and information technology plays the biggest role in achieving smart buildings. Despite the many benefits of smart buildings, there are disadvantages that can be attributed to a lack of financial resources, confidence in conducting new and untested technologies, professional ability to integrate and manage smart technologies, developers and owners' knowledge of the environmental impact of ineffective buildings. Several challenges like high costs of implementation and maintenance, security and privacy of data, behavioral inhibitions, difficulty in adaptation to the new environment and others exist. But, at a larger scale the plethora of advantages it offers and availability of solutions for these challenges makes smart building management a thing for the future. Finally, this research aims to implement in future a management system which controls and monitors the smart building and aids in achieving the energy efficiency sustainable and green energy of the current building, while maintaining the safety and security of occupants in the buildings.

## References

1. Rahiman R, Yenneti K, Panda A (2019) Making Indian cities energy smart. The Energy and Resources Institute (TERI), pp 1–24
2. Papantoniou S, Mangili S, Mangialenti I (2017) Using intelligent building energy management system for the integration of several systems to one overall monitoring and management system. *Energ Procedia* 111:639–647
3. Apanaviciene R, Vanagas A, Fokaide P (2020) Smart Building integration into a smart city (SBISC): development of a new evaluation framework. *MDPI* 13(9):1–19
4. Popescu D, Prada M (2013) Some aspects about smart building management systems-solutions for green, secure and smart buildings. *Re Adv Environ Sci* 126–132
5. Iwayemi A, Wan W, Zhou C (2011) Energy management for intelligent buildings. InTech
6. Vahidabegum UM, Rashmi NM, Rajashekar T, Pampapathi U, Kumuda B, Deepa B (2017) Renewable energy in smart buildings by managing green charge. *Int J Eng Sci Comput* 7(6):13101–13103
7. Abid M, Lghoul R, Benhaddou D (2017) ICT for renewable energy integration into smart buildings: IoT and big data approach. In: *IEEE Africon 2017 proceeding*. Researchgate, pp 883–888



8. Pan J, Jain R, Paul S, Vu T, Saifullah A, Sha M (2015) An Internet of Things framework for smart energy in buildings: designs, prototype, and experiments. *IEEE Internet Things J* 1–10
9. Eseosa O, Temitope F (2019) Review of smart based building management system. *World J Innov Res* 7(2):14–23
10. Basnayake J, Amarasinghe R, Attalage R, Udayanga T, Jayasekara A (2015) Artificial intelligence based smart building automation controller for energy efficiency improvements in existing buildings. *Int J Adv Inform Sci Technol* 40(40):85–91
11. Artificial intelligence in the field of building automation. [https://www.enocean-alliance.org/wp-content/uploads/2020/10/201004\\_Artificial-Intelligence-in-the-field-of-Building-Automation\\_final.pdf](https://www.enocean-alliance.org/wp-content/uploads/2020/10/201004_Artificial-Intelligence-in-the-field-of-Building-Automation_final.pdf), 25/5/2021
12. Luckey D, Fritz H, Legatiuk D, Dragos K, Smarsly K (2020) Artificial intelligence techniques for smart city applications. Researchgate, pp 1–15
13. ACEEE. <https://www.aceee.org/research-report/a1701>, 25/5/2021
14. Digital transformation monitor. <https://ec.europa.eu/growth/tools-databases/dem/monitor/tags/energy-efficiency>, 25/5/2021
15. Metallidou C, Psannis K, Egyptiadou E (2020) Energy efficiency in smart buildings: IoT approaches. *IEEE Access* 8:63679–63699
16. The guide to smart buildings. <https://www.trueoccupancy.com/smart-building-guide>, 25/5/2021
17. Sciuto D, Nacci AA (2014) On how to design smart energy-efficient buildings. In: The 12th IEEE international conference on embedded and ubiquitous computing (EUC). Researchgate, pp 1–5
18. Nacci AA et al (2014) A system C-based framework for the simulation of appliances networks in energy-aware smart spaces. In: IEEE world forum on Internet of Things. Researchgate, pp 1–6
19. Morvaj B, Lugaric L, Krajcar S (2011) Demonstrating smart buildings and smart grid features in a smart energy city. In: 3rd international youth conference on emergaics. Researchgate, Portugal, pp 1–9
20. Noack D et al (2011) Challenges and solution approaches for the online simulation of semiconductor wafer fabs. In: Proceedings of the winter simulation conference. Researchgate, pp 1845–1856
21. Rocha P, Siddiqui A, Stadler M (2015) Improving energy efficiency via smart building energy management systems: a comparison with policy measure. *Energ Build* 88(2015):203–213
22. Groissböck M, Heydari S, Mera A, Perea E, Siddiqui A, Stadler M (2014) Optimizing building energy operations via dynamic zonal temperature settings. *J Energ Eng* 140(1)
23. Cano EL, Moguerza JM, Ermolieva T, Ermoliev Y (2014) Energy efficiency and risk management in public buildings: strategic model for robust planning. *Comput Manag Sci* 11(1):25–44
24. Werner A, Uggen KT, Fodstad M, Lium A-G, Egging R (2014) Stochastic mixed integer programming for integrated portfolio planning in the LNG supply chain. *Energ J* 35 (1):79–97
25. Liang Y, Levine D, Shen Z-J (2012) Thermostats for the SmartGrid: models, benchmarks, and insight. *Energ J* 33(4):61–96
26. Skouby KE, Lynggaard P, Windekilde I, Henten A (2014) How IoT, AAI can contribute to smart home and smart cities services—the role of innovation. In: 25th European regional conference of the international telecommunications society. EconStor, Brussels, Belgium, pp1–13
27. Park H, Rhee B (2018) IoT-based smart building environment service for occupants' thermal comfort. *J Sens* 2018:1–10
28. Plagerasa P, Psannisa KE, Stergioua C, Wangb H, Guptac BB (2018) Efficient IoT-based sensor big data collection—processing and analysis in smart buildings. *Fut Gener Comput Syst* 82:349–357
29. Romdahani I, Al-Dubai A, YQasem M, Ghaleb B, Wadhaj I (2016) Cooja Simulator Manual. Technical Report, Edinburgh Napier University

30. Junior J, Souza M (2017) Barriers and challenges to smart buildings' concepts and technologies in Brazilian social housing projects. *Int J Sustain Real Estate Constr Econ* 1(1):31–50
31. Daissaoui A, Boulmakoul A, Karim L, Lbath A (2020) IoT and big data analytics for smart buildings: a survey. *Procedia Comput Sci* 170:161–168

# Optimal Packet Routing with Multiple Demands and Link Lines



Monu and L. N. Das

**Abstract** The information communication process uses circuit exchanging, to transmit information in a wired network. The LAN and WAN uses both circuit exchanging and datagram or virtual packet exchanging system to transfer analog–digital information through the computer processors. The Email, Fax message and internet digital information are processed through the analogue–digital network by circuit exchanging and transmission stage, it require packet exchanging in the datagram or virtual circuit exchanging to storage and transfer information in packet forms. This is done by the device named router and the router performs packet routing. During packets routing either the packet or blocks of information channelized through a fixed bandwidth allotted route, or a flexible bandwidth allotted route in the network. The nodes of the network are switches having storage capacity and the links are lines with flexible bandwidth capacity transfers the routed packets as per the demands of the nearest storage nodes or destination stations Thus the packet routing in a subscribed line-node network associates with costs. In this paper, we formulated a mathematical program similar with the problem mentioned in Ref. [1] to minimize the routing cost with considering all demanded routes having the known bandwidth capacity and different link line costs. The computation algorithm program is written in Python and a numerical example is also discussed to verify the codes.

**Keywords** Virtual Packet exchanging · Circuit exchanging · Minimum routing cost with fulfilling demands and capacity of the storage nodes and transfer link lines

## 1 Introduction

The information communication process uses circuit exchanging, to transmit information in a wired network. The LAN and WAN uses both circuit exchanging and datagram or virtual packet exchanging system to transfer analog–digital information

---

Monu (✉) · L. N. Das

Department of Applied Mathematics, Delhi Technological University, New Delhi, India

L. N. Das

e-mail: [lnidas@dce.ac.in](mailto:lnidas@dce.ac.in)

through the computer processors. The Email, Fax message and internet digital information are processed through the analogue–digital network by circuit exchanging and transmission stage, it require packet exchanging in the datagram or virtual circuit exchanging to storage and transfer information in packet forms. This is done by the device named router and the router performs packet routing. During packets routing either the packet or blocks of information channelized through a fixed bandwidth allotted route, or a flexible bandwidth allotted route in the network. The nodes of the network are switches having storage capacity and the links are lines with flexible bandwidth capacity transfers the routed packets as per the demands of the nearest storage nodes or destination stations Thus the packet routing in a subscribed line-node network associates with costs. Different software developmental efforts are made to solve the packet routing problem. In the works of Ref. [2] it is discussed certain issue in routing mechanism of packet forwarding. Prior to this study in the work of Refs. [3, 4], the random procedures and directory procedures information routing in communication networks are discussed. In the work of Ref. [5] the authors have discussed the packet routing and job-scheduling in congestion and dilation steps. In this paper we formulated a mathematical program similar with the problem mentioned in Ref. [1] to minimize the routing cost with considering all demanded routes having the known bandwidth capacity and different link line costs. The computation algorithm program is written in Python and a numerical example is also discussed to verify the codes.

Email, Fax message, internet digital information transmitter network uses router facilitated with packet exchanging, circuit exchanging, and message exchanging devices. The difference between circuit exchanging and packet exchanging is based on the types of connection. A packet exchanging is usually a connectionless entity without the physical path in the packet. Unlike circuit exchanging, packet exchanging does not require the use of a dedicated chance (Table 1).

## 2 Motivation

The packet exchanging mode of information transfer, transmits information across the digital network by breaking the information into blocks or packets for efficiently transfer the information by using various network devices. The network devices can then route the packets to the destination node or station, where the receiving device reassembles or sequences the packets for the use. Routing and exchanging are various functions of network communication. The function of exchanging is to switch the information packets between gadgets on the same network of the same LAN. The function of routing is to route packets between various networks or LANs. Any node in the packet exchanging network, can store the messages till the network transmission facility is available. Broadcast process of the packet exchanging mode of information transfer, requires much less bandwidth than the circuit exchanging message transmission, if unlimited sizes can be sent.

**Table 1** Distinction between circuit exchanging and packet exchanging

Characteristic of circuit exchanging	Characteristic of packet exchanging
In the circuit exchanging, there are three stages of actions namely connection establish, information transfer, and connection release	The packet exchanging directly transfer data, using a virtual circuit exchanging, if source destination address and packet sequence number is known
In the circuit exchanging mode of information transfer, every information unit realizes the whole fix address which is given by the source connection of the organization	In packet exchanging information transfer mode, every information unit simply realizes the last objective location moderate way is chosen by the switches
In the circuit exchanging mode of information transfer, information is processed at the source system only	In the packet exchanging form of information transfer, information is processed at all intermediate node controlled by the source system
Resource reservation is the feature of the circuit exchanging mode of information transfer, because the path is fixed for information transmitting	There is no resource reservation features in a packet exchanging mode of information transfer, because the bandwidth is shared among the users
The information transfer through the circuit exchanging is more solid	Packet exchanging is less solid
Information are not stored and then forwarded at the stages in the circuit exchanging	Information are stored then forwarded in the packet exchanging technique
Transmission of the information is finished by the source framework	Transmission of the information is done not only by the source computing device, but also by the peripheral routers
In the circuit exchanging mode of information transfer, a congestion can occur during connection establishment time, there are cases for requesting a channel of the channel is already occupied	In the packet exchanging information transfer form a congestion can occur during the transfer phase, as the large numbers or packets comes in no time
Circuit exchanging mode of information transfer is not advantageous for respective traffic	The packet exchanging mode of information transfer is suitable for handling respective traffic
In-circuit exchanging, mode of information transfer, electric charge relies upon time and distance not in rush hour gridlock in the network	In packet exchanging, mode of information transfer, the charge is passed on the number of bytes and connects with the times
Recording of the packet is never conceivable on the circuit switch method of information move	While the recording of the packet is possible in the packet exchanging mode of information transfer

The internet exchanging protocol consists of Frame relay, multi-protocol label exchanging, Ethernet, Asynchronous transfer mode, IPv4, spanning tree protocol, Address resolution protocol or (ARP), Internet control message protocol (ICMP).

There are many types of exchanging namely packet, circuit, multi-layer, virtual-circuit, wide areas network (WAN), and LAN, circuiting virtual circuit exchanging always refer to WAN or telephone technology packet exchanging usually concerns a

router or perhaps a WAN switch. Multi-layer exchanging is a technique for improving the processing of IP packets.

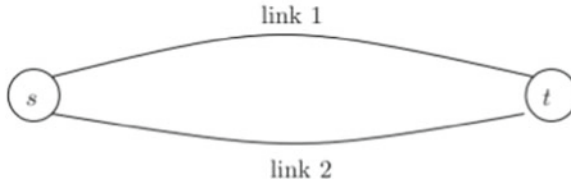
Packet exchanging routes either on a fixed bandwidth allowed route or a flexible bandwidth allotted route in the network. The packet routing switches are either the datagram exchanging or virtual circuit exchanging. The structure of the datagram packet consists of a header, information package, and trailer. The header consists of source and destination addresses and trailer consists of sequence numbers and transmission time digital information. The flexible bandwidth routing is also named flooding or selective node flooding. The optimal selection of nodes and links in the path of a network having definite hops connecting multi-lines with known channel capacity bandwidth allotment is an optimal routing problem, which is solved in Sect. 3.2. One among the many algorithms to select the effective lines and nodes in the routing network selection process is known as Dijakstra Algorithm. In this algorithm, the effective minimum cost path is computed by the system either by a central table that contains the rows and columns having link costs or by the distributed tables with link cost specifications.

### 3 Problem Description

The problem we are describing in this paper is an optimal packet routing problem mentioned in [1]. It consists of  $m$  information packets flows through the routed network from a source node  $s$  to a destination node  $t$  following one of the two possible links. Link1 and Link 2 capacities  $u_1 = 1$  and  $u_2 = 2$  respectively. The packet routing through the link 2 is 30 more expensive than the packet routing through the link 1 (Table 2).

**Table 2** Demand of the route, required capacity and the cost on links specification

Demand frequency	Required capacity Mbps	Cost on link 1	Cost on link 2
1	0.4	200	250
2	0.3	200	225
3	0.2	150	200
4	0.1	100	150
5	0.3	200	225
6	0.3	200	225
7	0.5	400	500
8	0.2	150	200
9	0.2	150	200
10	0.6	500	550



Formulated an optimization problem which is to minimize the packet routing cost on choosing a suitable route. The mathematical program model would be acceptable to generalize the situation with  $m$  possible parallel links between  $s$  and  $t$ ?

### 3.1 Definition and Notation

The following definition and notations are used in the mathematical formulation of the problem.

- Indicators:
- $i$ : list on the arrangement of requests  $M = (1, 2, \dots, n)$ .
- $j$ : list on the arrangement of connections  $L = (1, 2, \dots, m)$ .
- Parameters:
- $a_i$ : limit (capacity) utilized by requests  $i$ ;
- $c_i$ : routing cost for demand  $i$  on link 1;
- $p_j$ : cost rate contrast between on interface  $j$  and connection 1;
- $u_j$ : installed limit of the capacity on interface  $j$ .
- Variables:

$X_{ij} = 1$  if packet  $i$  is directed on interface  $j$ , 0 otherwise.

- Objective function:  
Min

$$\sum_{j=1}^m \sum_{i=1}^n (p_j + 1) c_j x_{ij}$$

- Constraints:

$$\sum_{j=1}^m x_{ij} = 1 \quad (1)$$

$$\sum_{i=1}^n a_i x_{ij} \leq u_j \quad (2)$$

### 3.2 *Computation Results*

See Fig. 1.

### 3.3 *Numerical Example*

In the example, we take a set of three demands and three links i.e. 1, 2, 3. The value of  $p_1 = 1$ ,  $p_2 = 2$ ,  $p_3 = 3$  and  $u_1 = 2$ ,  $u_2 = 3$ ,  $u_3 = 4$ . The Optimal solution is  $X_{11} = X_{21} = X_{31} = 1$  and other remaining are zero, and the total cost is 30 (Fig. 2).

## 4 Conclusion

The characteristics of our presented algorithm and computation codes we have specified is compared with the characteristic of Dijkstra Algorithm, which is briefed in Table 3.



```

In [ ]:

from pulp import *
import pandas as pd
import numpy as np

#set of demands, i = (1,2,...,n)
n_demands = int(input("Enter the value of demands = "))

#set of links, j = (1,2,...,m)
n_links = int(input("Enter the values of links = "))

# Routing cost for demand i on link l
row = int(input("Enter the number of rows:"))
column = int(input("Enter the number of columns:"))

# Initialize empty matrix
matrix = []
print("Enter the entries row wise:")

# For user input
for i in range(row):          # A outer for loop for row entries
    a = []
    for j in range(column):    # A inner for loop for column entries
        a.append(int(input()))
    matrix.append(a)

# For printing the matrix
for i in range(row):
    for j in range(column):
        print(matrix[i][j], end = " ")
    print()
A_row = int(input("Enter the value of row of a :"))
A_column = int(input("Enter the value of column of a :"))
a = np.random.random((A_row,A_column))
a
print(a)

model = LpProblem("Problem", LpMinimize)
variable_names = [str(i)+str(j) for j in range(1, n_links+1) for i in range(1, n_demands+1)]
variable_names.sort()
print("Variable Indices:", variable_names)

DV_variables = LpVariable.matrix("X", variable_names, cat = "Integer", lowBound= 0 )
allocation = np.array(DV_variables).reshape(n_demands,n_links)
print("Decision Variable/Allocation Matrix: ")
print(allocation)
#p_1 = int(input(Enter the value of p_1:))
#p = np.array([[2],[3],[4]])
#p
#print(p)

R = int(input("Enter the number of rows of p:"))
C = int(input("Enter the number of columns p:"))

print("Enter the entries in a single line (separated by space): ")

# User input of entries in a
# single line separated by space
entries = list(map(int, input().split()))

# For printing the matrix
p_matrix = np.array(entries).reshape(R, C)
print(p_matrix)

res = p_matrix*allocation
print(res)

L = int(input("Enter the number of rows of u:"))
M = int(input("Enter the number of columns u:"))

```

**Fig. 1** Python code of problem described in paper

```

print("Enter the entries in a single line (separated by space): ")

# User input of entries in a
# single line separated by space
entries = list(map(int, input().split()))

# For printing the matrix
u_matrix = np.array(entries).reshape(L, M)
print(u_matrix)

obj_func = lpSum(res*matrix)
print(obj_func)
model += obj_func
print(model)

# First constraints:
for i in range (n_demands):
    print(lpSum(allocation[i][j] for j in range(n_links)) == 1)
    model += lpSum(allocation[i][j] for j in range(n_links)) == 1

# Second Constraints:
for j in range(n_links):
    print(lpSum(a[j]*allocation[i][j] for i in range(n_demands)) <= u_matrix[j])
    model += lpSum(a[j]*allocation[i][j] for i in range(n_demands)) <= u_matrix[j]

#model.solve()
model.solve(PULP_CBC_CMD())

status = LpStatus[model.status]

print(status)

print("Total Cost:", model.objective.value())

# Decision Variables
for v in model.variables():
    try:
        print(v.name,"=", v.value())
    except:
        print("error couldnt find value")

```

**Fig. 1** (continued)

## References

1. Liberti L Problems and exercises in operations research
2. Devikar RN, Patil DV, Chandraprakash V (2016) Issues in routing mechanism for packet forwarding: a survey. Int J Electr Comput Eng (IJECE) 6(1):421–430
3. Prosser RT (1962) Routing procedures in communication networks-Part I: random procedures. IRE Trans Commun Syst CS-10:322–329
4. Prosser RT (1962) Routing procedures in communications networks-Part II: directory procedures. IRE Trans Commun Syst 3–10:329–335
5. Leighton FT, Maggs BM, Rao SB (1994) Packet routing and job-scheduling in O(congestion+dilation) steps. Combinatorica 14:167–186

```

MINIMIZE
1*X_11 + 2*X_12 + 3*X_13 + 8*X_21 + 10*X_22 + 12*X_23 + 21*X_31 + 24*X_32 + 27*X_33 + 0
VARIABLES
0 <= X_11 Integer
0 <= X_12 Integer
0 <= X_13 Integer
0 <= X_21 Integer
0 <= X_22 Integer
0 <= X_23 Integer
0 <= X_31 Integer
0 <= X_32 Integer
0 <= X_33 Integer

X_11 + X_12 + X_13 = 1
X_21 + X_22 + X_23 = 1
X_31 + X_32 + X_33 = 1
0.06881885034551005*X_11 + 0.06881885034551005*X_21 + 0.06881885034551005*X_31 <= 2.0
0.12080423510017313*X_12 + 0.12080423510017313*X_22 + 0.12080423510017313*X_32 <= 3.0
0.11647681229982221*X_13 + 0.11647681229982221*X_23 + 0.11647681229982221*X_33 <= 4.0
Optimal
Total Cost: 30.0
X_11 = 1.0
X_12 = 0.0
X_13 = 0.0
X_21 = 1.0
X_22 = 0.0
X_23 = 0.0
X_31 = 1.0
X_32 = 0.0
X_33 = 0.0

```

**Fig. 2** Final of computation of example**Table 3** Difference between Dijkstra's algorithm and presented algorithm

Characteristic of Dijkstra's algorithm	Characteristic of our presented algorithm
It compute the least-cost path from one node to all other nodes in the network	We compute the cost of a path from one node to another Node with various demands
It is Iterative Algorithm	It is not an Iterative Algorithm
Each Node has complete information about the cost of all links	No Node has complete information about the cost of all links
Each Node have some finite cost and also links have finite cost	No Node have finite cost but links have finite cost
Each link begin with the knowledge of the cost of the Nodes and the knowledge of the cost of its own link directly attached between the Nodes	Each link begin with only the knowledge of the cost of its own directly attached between the Nodes, known demand lines with known capacity

# Mapping and Change Detection of Mangroves Using Remote Sensing and Google Earth Engine: A Case Study



Geetanjali S. Mahamunkar, Arvind W. Kiwelekar, and Laxman D. Netak

**Abstract** In this case study, we describe the mapping and change detection in the mangrove coverage of Raigad District of Maharashtra, India, using remotely sensed images. The mangrove maps are prepared using the Landsat8 and Landsat7 satellite data for 2019 and 2009, respectively. To develop machine learning models, we use a public domain dataset for sampling mangroves from these Landsat images. We apply the random forest algorithm to observe the changes in the mangrove coverage. The purpose of this case study is to perform temporal analysis on spatial data and also to evaluate the applicability of Google Earth Engine for implementing change detection in mangrove coverage.

**Keywords** Remote sensing · Mangroves · Change detection · Spatial–temporal analysis

## 1 Introduction

The ecological importance of mangroves, such as coastal defence, fishery support, climate control, livelihood, wood for fuel and other ecosystem services, necessitates the development of conservation strategies. Land surveys are time-consuming and expensive when it comes to mapping mangroves. The mangrove coverage is constantly changing as a result of urbanization and anthropogenic activities. Remote sensing is a commonly used method for observing and collecting data about the Earth without having to make direct contact with it [1]. It can be used to assess the biomass of mangroves as well as their ecosystem status, i.e. whether they are

---

G. S. Mahamunkar (✉) · A. W. Kiwelekar · L. D. Netak  
Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University,  
Lonere, Raigad 402103, India  
e-mail: [gsmahamunkar@dbatu.ac.in](mailto:gsmahamunkar@dbatu.ac.in)

A. W. Kiwelekar  
e-mail: [awk@dbatu.ac.in](mailto:awk@dbatu.ac.in)

L. D. Netak  
e-mail: [ldnetak@dbatu.ac.in](mailto:ldnetak@dbatu.ac.in)

intact or degraded. This will aid in the global management and regeneration of mangrove forests. The observation of a major transition in a location on Earth between two points in time is known as change detection. The change detection in remotely sensed images [2] is useful for various applications such as (i) monitoring urban green space changes [3], (ii) keeping a check on urban encroachment, (iii) deforestation, and (iv) land cover and land use changes [4–7]. Identifying the geographical location and types of changes that occur, quantifying those changes, and determining the accuracy of the findings are some of the objectives of change detection. The spectral and spatial resolution of images produced by a remote sensor is a critical consideration. Spatial resolution refers to the amount of physical area on Earth that corresponds to a pixel in a raster picture. The number of electromagnetic bands used in a pixel is referred to as spectral resolution, which usually corresponds to three bands to seven visible colours. The changes in spectral values of pixels of remotely sensed images can be helpful in change detection as these images are available at various resolution levels and thus make a vital data source for this purpose. However, various challenges are faced while using remotely sensed images. For example, normally, remotely sensed images are collected using various satellites and multi-spectral sensors, which are influenced by atmospheric interference; there is also a difficulty in assessing and monitoring mangroves using remote sensing since they are mostly located in tropical coastal areas, which are typically cloudy and therefore obscure vision in satellite images. As a result, these images must be improved before being analysed for change detection. Different land cover forms can be classified using image processing software. For accurate change detection, the right image processing method is crucial. It should be selected based on the image type, sensor orientation, and type of image noise. We used Google Earth Engine in this case study because it incorporates various sources of knowledge into change detection. Its ability to provide a wider view of the region under investigation as well as regular coverage of the area is a plus. Change detection will be simpler with this technology, and it will be more accurate and less expensive. Due to the regular availability of Landsat data and the imagery's medium spatial resolution, Landsat imagery is primarily used to detect changes to the Earth's surface. As a result, rather than relying on traditional methods, a spatial–temporal analysis using Geographic Information System technology should be carried out [8].

## 2 Literature Review

Spectral bands, spatial resolution, and temporal details are all included in satellite images. The spectral vegetation indices can be calculated using the different spectral bands associated with the dataset. Vegetation indices are a ratio of the different bands in satellite images. For vegetation analysis, there are a variety of such indices [9, 10]. In this study, we have used spectral indices such as the normalized difference vegetation index (NDVI), the normalized difference mangrove index (NDMI), the modified normalized difference water index (MNDWI), the simple ratio (SR), the

band ratio 54, the band ratio 35, and the green chlorophyll vegetation index (GCVI) to distinguish mangroves from other vegetation in the images obtained for this analysis.

However for more accurate identification of mangroves by creating the training and testing samples for the algorithm, we have used the mangrove dataset [11] along with the spectral indices. The random forest algorithm is used as the classification algorithm in this analysis since it has many advantages over other image classification methods. It is nonparametric, can handle both continuous and categorical data, is easy to parameterize, resists over-fitting, can handle outliers in training data, and tests ancillary data like classification error and variable significance [12].

### 3 Study Area

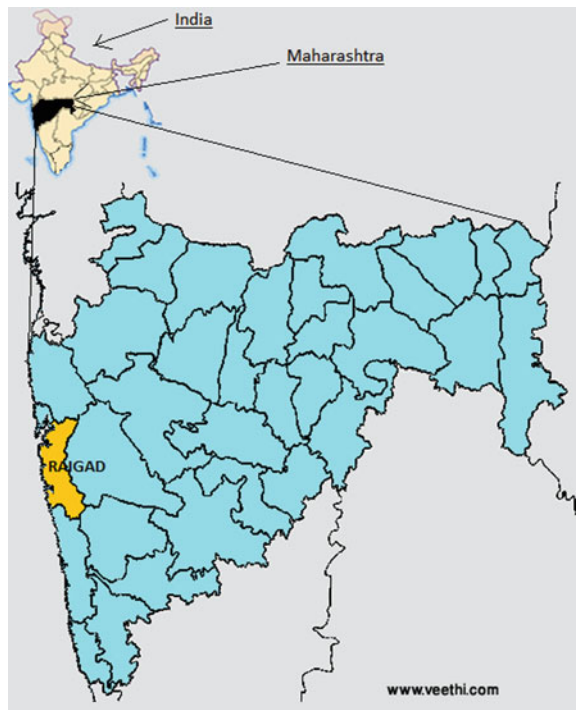
In this case study, we have calculated the mangrove coverage area in square kilometers for the Raigad District in the state of Maharashtra, India. Raigad is a district close to Mumbai with an area of 7152 km<sup>2</sup>. Mangroves can be found in seven coastal districts of Maharashtra [13]. According to [14], Raigad has the most mangrove coverage in Maharashtra, accounting for around 40% of the state's total mangrove coverage. The current study uses Google Earth Engine to map the area covered by mangroves in the coastal area of Raigad District, as well as detect changes in mangrove coverage over time. Figure 1 shows the location of the study area in India taken from [15].

### 4 Methodology

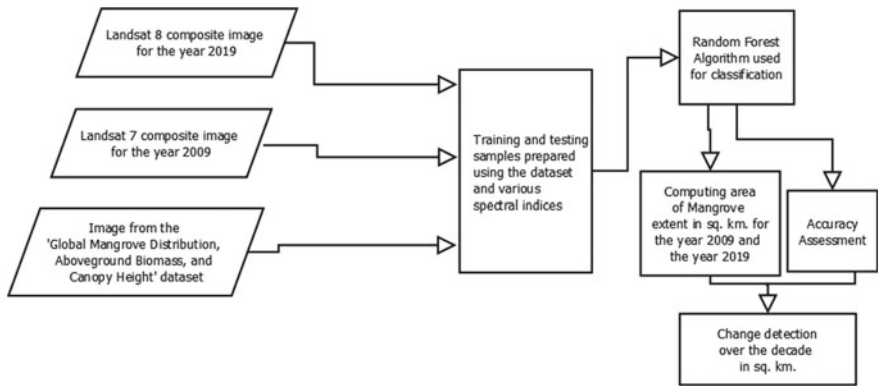
In this case study, the change detection in mangrove coverage is determined using the methodology shown in Fig. 2.

#### 4.1 *Random Forest Classification*

Random Forest classification, an ensemble learning process, is used in this case study. Ensemble learning is a machine learning paradigm that involves training several models to solve a common problem. A random forest is a collection of multiple decision trees in which a random set of features is used to split on at each node. Instead of using all of the input variables, the best split of the decision tree is selected at each node from a random sample of them. The final decision is made based on the predictions from all of the trees: mode of classes for classification and mean prediction for regression.



**Fig. 1** Location of Raigad District in Maharashtra, India



**Fig. 2** Methodology to determine change in mangrove coverage using random forest algorithm in Google Earth Engine

## 4.2 Google Earth Engine (GEE) for Implementing Random Forest Algorithm

Though vegetation indices alone are adequate for detecting changes in landcover, supervised machine learning algorithms such as SVM, random forest, and naive Bayes classifiers can be used to improve accuracy in detecting and analysing changes [12, 16]. The classifier package of Earth Engine handles supervised classification by traditional machine learning algorithms [17]. These classifiers include CART, random forest, naive Bayes and SVM. Objects and methods make up the Earth Engine API. Raster images, vector features, numbers, and strings are all represented by objects. Each of these objects is a member of a particular class, and each class has a limited set of functions. Workflow scripts combine objects and methods and send them to Earth Engine servers for processing. Some of the common Earth Engine object classes which we have used in our implementation are *Image*, *ImageCollection*, *Geometry*, *Feature* and *FeatureCollection* where *Image* is a fundamental raster datatype, *ImageCollection* is a set of images, *Geometry* is the fundamental vector data type in Earth Engine, *Feature* is a geometry with attributes and *FeatureCollection* is set of features.

The steps followed for classification in GEE are: The first step is to gather training data, which is a *FeatureCollection* with a class label storing property and properties for predictor variables. The predictors should be numeric and the class labels should be sequential with integers starting at 0. Data for training and/or validation may come from a number of sources. It can be collected interactively in Earth Engine using the geometry drawing tools. Alternatively, predefined training data can be imported from an Earth Engine Asset. In this case study, we have used both. The public domain dataset [11] has been imported from the Earth Engine Asset.

The next step is to obtain a classifier using one of the constructors in `ee.Classifier`. This classifier is trained using `train()` function and then used to classify an image or *FeatureCollection* using `classify()` function.

Finally, using independent validation data, calculate the classification error. To access the accuracy of a classifier *ConfusionMatrix* can be used. In this case study, we have created each sample of the size same as that of the Landsat pixel as data should be of same spatial resolution to perform classification. The samples are then split into training and testing dataset using `randomColumn()`.

## 4.3 Preprocessing

To prevent misclassification, a multispectral satellite image must be pre-processed before implementing any algorithm. A 'pixel\_qa' band in Landsat data can be used to construct a cloud masking function. Cloud masking is performed on the satellite images using the following function prior to classification in the present study:



**maskClouds(image):**

1. Cloud shadow and cloud are represented by bits 3 and 5, respectively.
2. Get the pixel QA band from the Landsat image.
3. Perform *bitwiseand* operation between pixel QA and cloud shadow bit mask as well as pixel QA and cloud bit mask to set both flags to zero and indicate clear conditions.
4. Return the masked image, scaled to [0, 1].

To be able to distinguish vegetated regions in Landsat imagery, various spectral indices are applied. The image is masked for clouds and cloud shadows after adding temporal parameters to it. Later a composite image is created on a per pixel, per-band basis using a median Reducer *median()* and clipped to the area of interest. We clipped the SRTM data to our region of interest and built an elevation mask since mangroves can reach a maximum height of 65 m. This mask, along with the masks generated using the NDVI and MNDWI bands, is then added to the composite image. This enables us to create Landsat composite images for the years 2019 and 2009, respectively, using Landsat8 and Landsat7. This case study uses the assumption that mangrove (M) is a collection and that different vegetations (x) are members of the region under consideration to decide if the vegetation is mangrove or not.

$$x \in M \quad (1)$$

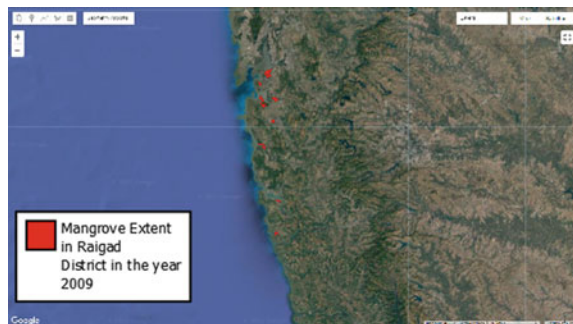
In Formula 1, x can be a point or a polygon, which is the type of geometry Earth Engine uses to handle vector data. To implement the random forest algorithm, this case study follows all of the steps outlined in Sect. 4.2. The model is run with 100 trees and 5 randomly selected predictors per split using the classifier *.smileRandomForest*.

The classifier performs the following steps for the same:

- Train by selecting the desired bands and landcover property
- Retrieve the landcover property from classes

At last, an image is created to mask the result to display only mangrove extent as shown in Figs. 3 and 4.

**Fig. 3** Mangrove extent of Raigad District in the year 2009



**Fig. 4** Mangrove extent of Raigad District in the year 2019



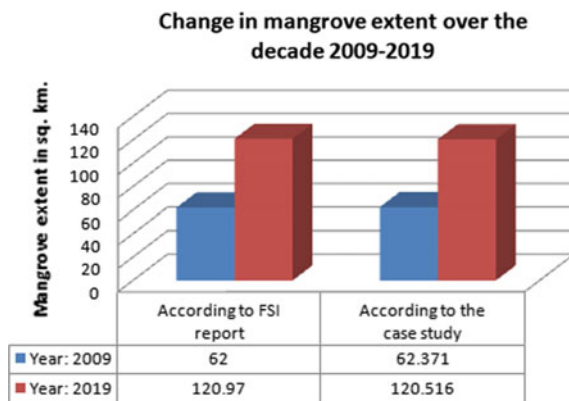
5 Result

The 30 m resolution Landsat 7 and Landsat 8 datasets were used to map mangroves in 2009 and 2019, respectively. The Landsat 7 dataset covers the period from 1999 to the present, while Landsat 8 covers the period from 2013 to the present. These datasets also provide data that has been atmospherically corrected for the mentioned time span. This study relies on the use of the above datasets since it seeks to examine changes in mangroves for the Raigad District over the last decade, from 2009 to 2019. Table 1 summarizes the output of the study.

**Table 1** Result of the case study

Year:	2019	2009
Total number of samples (N):	184	182
Number of training samples out of N:	129	117
Number of testing samples out of N:	55	65
Validation overall accuracy percent:	94.54	89.23
Training overall accuracy percent:	98.44	99.14
Kappa coefficient:	0.877	0.784
Mangrove extent in km <sup>2</sup> calculated by the case study:	120.516	62.371
Actual Mangrove extent in km <sup>2</sup> according to [14]:	120.97	62.0

**Fig. 5** Comparison of result of the case study with FSI-2019 report data



## 6 Conclusion and Future Scope

The overall accuracy calculated from the training data is significantly higher than that calculated from the validation data. Since the random forest is ‘fit’ to the training data, the accuracy estimated from training data is an overestimate. As the estimate from the validation data shows that the estimated accuracy on unknown data is lower. There is increase in mangrove coverage by 58.97 km<sup>2</sup> over the decade from 2009 to 2019 as per [14], and the result obtained by our case study is 58.145 km<sup>2</sup> as shown in Fig. 5 which is approximately the same. Thus, using remote sensing and Google Earth Engine to estimate mangrove coverage and change over time is more effective than conducting ground surveys. We can reach inaccessible areas where mangroves are present using remote sensing. Using Geographic Information System technology like Google Earth Engine, temporal–spatial analysis can be made simpler. Deep learning algorithms may be used to perform spatial–temporal analysis instead of machine learning algorithms, resulting in even higher accuracy. These methods may be used by government agencies responsible for the protection of mangroves to monitor their changes.

## References

1. Kiwelekar AW, Mahamunkar GS, Netak LD, Nikam VB (2020) Deep learning techniques for geospatial data analysis. In: Machine learning paradigms. Springer, pp. 63–81
2. Asokan A, Anitha J (2019) Change detection techniques for remote sensing applications: a survey. *Earth Sci Inf* 12(2):143–160
3. More N, Nikam VB, Banerjee B (2020) Machine learning on high performance computing for urban greenspace change detection: satellite image data fusion approach. *Int J Image Data Fusion* 1–15

4. Haque MdI, Basak R (2017) Land cover change detection using gis and remote sensing techniques: a spatio-temporal study on tanguar haor, sunamganj, bangladesh. *The Egypt J Rem Sens Space Sci* 20(2):251–263
5. Chunlei M, Bin A, Jun Z, Xu X, Wei H (2019) Change detection of mangrove forests in coastal guangdong during the past three decades based on remote sensing data. *Remote Sens* 11(8):921
6. Pham L, Brabyn L (2017) Monitoring mangrove biomass change in vietnam using spot images and an object-based approach combined with machine learning algorithms. *ISPRS J Photogrammetry Rem Sens* 128:86–97
7. Jianing Z, Jinguan L, Guozhuang S (2018) Mapping mangrove forests of dongzhaigang nature reserve in china using landsat 8 and radarsat-2 polarimetric sar data. *Sensors* 18(11):4012
8. Cárdenas NY, Joyce KE, Maier SW (2017) Monitoring mangrove forests: are we taking full advantage of technology? *Int J Appl Earth Obs Geoinformation* 63:1–14
9. Chuvieco E, Huete A (2009) Fundamentals of satellite remote sensing
10. Gupta K, Mukhopadhyay A, Giri S, Chanda A, Datta Majumdar S, Samanta S, Mitra D, Samal RN, Pattnaik AK, Hazra S (2018) An index for discrimination of mangroves from non-mangroves using landsat 8 oli imagery. *MethodsX* 5:1129–1139
11. Simard M, Fatoyinbo T, Smetanka CV, Rivera-Monroy H, Castaneda-Mova E, Thomas N, Van Der Stocken T (2019) Aboveground Biomass, and Canopy Height, Global Mangrove Distribution
12. Horning N et al (2010) Random forests: an algorithm for image classification and generation of continuous fields data sets. In: *Proceedings of the International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences*, Osaka, Japan, volume 911 (2010)
13. <https://mangroves.maharashtra.gov.in/Site/1120/Mangroves%20of%20Maharashtra>
14. FSI-ISFR (2019) Table: Forest cover change in Maharashtra (Area in sq.km.). Forest Survey of India (Ministry of Environment Forest and Climate Change) Kaulagarh road, P.O. IPE Dehradun - 248195, Uttarakhand, India
15. <http://www.veethi.com>
16. Heumann BW (2011) An object-based classification of mangroves using a hybrid decision tree-support vector machine approach. *Remote Sens* 3(11):2440–2460
17. <https://developers.google.com/earth-engine/guides>

# Source Code Obfuscation: Novel Technique and Implementation



Krishna Yadav, Rutuja Kamble, Sumit Kale, Yash Pazare, and Pallavi Ahire

**Abstract** In today's world the sheer amount and increasing range of unlicensed practices by injecting malicious programs, hackers are able to gain forbidden access to the devices and private information. Fortunately, the businesses who frame up the web world and therefore the security professionals that support them have their own set of repositories to retaliate. The basic idea behind source code obfuscation is to alter the computer programs in order to cover their delicate information while retaining their original functionality. The objective of the proposed method is to keep the code basic, without any crucial changes, here the invader will be unable to recognize the difference between the original code and the obfuscated code, and also, the code will be well obfuscated. Results evaluate that to what scale the code obfuscation is in a position to form attacks harder to be performed and disclose that obfuscation can attenuate the result of things which will alter the odds of a successful attack.

**Keywords** Code obfuscation · Decryption · Encryption · Functionality

## 1 Introduction

Sensitive data such as license numbers or deciphered keys are usually present in our software programs which could be the prior target of an attacker just in case the code is scanned or examined by malicious user's attack model. Among the feasible protection ways which will be applied to limit malicious program, code obfuscation aims at complicating the illustration of code. The original values are highly secured using program transformation. Basically, it is known as code obfuscation.

The possibility to hide constant values could also be a basic building block of many obfuscation techniques. Secondly, this conjointly contributes a protection against

---

K. Yadav (✉) · R. Kamble · S. Kale · Y. Pazare  
Sinhgad Institute of Technology, Lonavla, India

P. Ahire  
Department of Technology, Savitribai Phule Pune University, Pune, India  
e-mail: [pallaviyahire@sinhgad.edu](mailto:pallaviyahire@sinhgad.edu)

malicious modifications to a program. This stems from the actual fact that an attacker first must acknowledge the program associate in creating nominal modifications. The main objective behind software protection through data obfuscation is to alter the source code of a software in disguised manner.

To protect data that is classified as in personal information, sensitive personal data, or commercially sensitive data. Data obfuscation represents a whole set of code protection techniques that raise the bar for the offender in terms of skills or resources. Obfuscation techniques are of general class code protections wide adopted to stop malicious meddling of the code by creating applications tougher to know and therefore tougher to change. Code and data obfuscation are two basic types of obfuscation techniques counting onto protected assets. Code Obfuscation is a code transformation technique through which practicality of original code is maintained whereas the obfuscated code is tough to reverse engineer. Obfuscation technique manipulates the ASCII text file to form it harder to research and harder to know for the attacker.

## 2 Background

Protection of a software has become progressively an important crucial demand for industrial software development. Every software provider should remember the possibility for man-at-the-end (MATE) attacks against their system and therefore the techniques available to attenuate these attacks. While introductory observed researches have been performed to work out the impact of code obfuscation, our prior task is to focus at a survey of obfuscation and then shift our focus to outlining various data obfuscation techniques.

In this presentation, we present one novel code obfuscation technique that is being executed on multiplication operator '\*' which may guide us to the different obfuscation areas.

### 2.1 *Related Work*

We have gone through the following mechanisms from the respective papers which guided us throughout the procedure. Along with the help of the respective citations, we marked the references given in the paper.

From the number of methods, we have shortlisted few of them by understanding their functionality and implementation process, and thorough study and analysis is done, which helped us to dig into the right path for further development of our research. The given methods also helped us to decipher the pros and cons of priorly developed obfuscated techniques (Table 1).

**Table 1** Literature review

Reference paper	Methodology	Observations
1 Mechanisms for Source Code Obfuscation in C: Novel Techniques and Implementation	<ul style="list-style-type: none"> <li>– <b>First method</b> is through reference pointer mechanism</li> <li>– <b>Second method</b> is using the shift operation to perform the data obfuscation. Bitwise operation is used for all the procedures</li> <li>– <b>Third method</b> is that three different bitwise operators are used to perform obfuscation (Left-shift, bitwise AND, and XOR-Shift)</li> <li>– <b>Fourth method</b> XOR-VarMerge, XOR masking obfuscation encrypts a transparent value of a variable by calculating its bitwise XOR (denoted by <math>\wedge</math>) with a mask (integer constant)</li> </ul>	<ul style="list-style-type: none"> <li>– The paper consists of seven C obfuscators and four data different techniques of data obfuscation</li> <li>– These techniques are used on ‘+’ addition operator which may lead us to a different obfuscation arena</li> <li>– Each of the four methods proposed works perfectly fine, the difference between their execution time is negligible, and no error is produced in the process</li> </ul>
2 Obscuring Mobile Agents by Source code obfuscation	<ul style="list-style-type: none"> <li>– The given paper shows the technique in which the implementation is done by parser that makes the given program as input and obfuscates it which produces an obfuscated code/program that has the same working as the original with negligible overheads</li> <li>– The program compilation is done, and the byte code is sent</li> </ul>	<ul style="list-style-type: none"> <li>– The output code appeared to be scrambled</li> <li>– Code obfuscation provides added security for a brief period of your time that is adequate for the mobile agent to finish the task and move on to next node</li> </ul>
3 Assessment of Data Obfuscation with Residue Number Coding	<ul style="list-style-type: none"> <li>– The idea of ‘residue number coding’ encrypts program variables to cover their real values, while supporting functions within the encrypted domain</li> <li>– Update value representation without substituting how values are stored in memory</li> <li>– Instead of being vulnerable to an attack by a hacker, it is better to have a costlier as well as a secure strategy which is code obfuscation</li> </ul>	<ul style="list-style-type: none"> <li>– Auxiliary computational overhead is needed when obfuscated variables need to be encrypted/decrypted whenever an operation is carried out that is not supported by the encoding strategy</li> </ul>

(continued)

## 2.2 Obfuscation Taxonomy

Obfuscation methods are approachable using various techniques, they are as follows:

**Table 1** (continued)

Reference paper	Methodology	Observations
4	Simple Obfuscation tools for software protection <ul style="list-style-type: none"> <li>– Here Java programming language and IntelliJ IDEA development environment was used. Also, Apache Maven was used for managing dependencies and external libraries for implementation</li> <li>– Implemented using JavaFX library, it is a group of graphics and media packages that permits the developers to design, create, test, debug, and deploy client applications</li> <li>– Also, external module functionality is put into use</li> </ul>	<ul style="list-style-type: none"> <li>– The tool results to perform both data flow and control flow obfuscation</li> <li>– It also provides GUI of better usability to easily configure the transformation process</li> <li>– It makes the application to operate consistently on diverse platforms</li> </ul>

The basic approaches of obfuscation are code-oriented and data-oriented which means that we can perform a particular kind of obfuscation methodology in either way. These two orientations are further subdivided.

- **Code-Oriented Obfuscation:** Code obfuscation is achieved by updating workable code so it is not any longer helpful to a hacker, however, remains totally useful. Whereas technique could modify the particular method instructions (meta-data), it does not alter the output of the program. Code obfuscation consists of the many completely different techniques which will complement one another to form a superimposed defense. It is simplest for languages that make some kind of intermediate level directions.
- **Data-Oriented Obfuscation:** Data obfuscation hides the data from the unauthorized user. The manner in which data are hidden can result in huge savings and reduces the risk of breaches and fine. This technique yields many benefits as well. These two orientations are further subdivided which are given below.

### 3 Proposed mechanisms

#### 3.1 *Multiplication of Two Numbers by Using Reference Pointer for Any Two-Positive Integer*

As proposed by Ahire [1] and Abraham [1], we introduced the technique named reference pointer mechanism for multiplication creating the code obfuscation. While selecting this technique, we mainly focused on the references and its functions. In Fig. 2, a pointer functionality is taken into consideration.

Operation on the original variables is masked using function inside the return type. All the advantages and disadvantages of this technique are taken into consideration. This technique can be implemented with various software programs written in C.



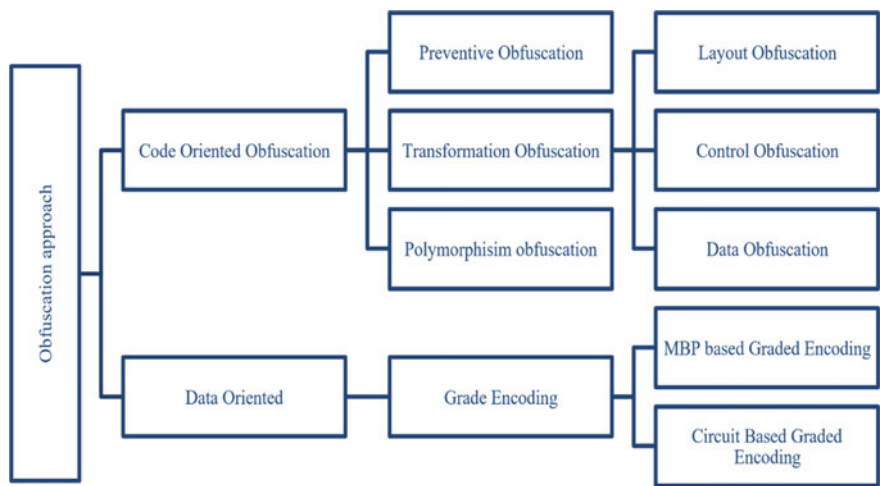


Fig. 1 Obfuscation taxonomy

Original .C	Obfuscated .C
<pre>#include &lt;stdio.h&gt; int main() {     int a, b;     printf ("Enter two numbers:");     scanf ("%d %d", &amp;a, &amp;b);     printf ("%d", a*b);     return 0; }</pre>	<pre>#include &lt;stdio.h&gt; int func(int a, int b) {     const char *k=0;     if ( y == 0 )         return (0);     return ( y &gt;= 1)? ( &amp;k[x] + func( &amp;k[x], &amp;k[y-1])):     ( &amp;k[-x] + func (&amp;k[x], &amp;k[y+1])); } int main() {     int a, b;     printf ("Enter two numbers");     scanf ("%d %d", &amp;a, &amp;b);     printf ("%d", func(a,b));     return 0; }</pre>

Fig. 2 Example of original and obfuscated code using reference pointer mechanism for multipli-  
cation

So, multiplication operator over here is obfuscated within the execution time pace equivalent to original code.

For further consideration, two columns are provided which eventually shows the difference between the original and the obfuscated code where by transforming few lines of the code irrespective of the negligible difference in the execution time, the entire skeleton and the functionality of the code remain the same without any loss to the authenticity of code (Fig. 2).

### Algorithm 1: Multiplication by Reference Pointer

```

Function: func (x, y)
// Input: Two integers x and y.
// Output: Multiplication of two integer number.
1. Begin
2. if y = 0, then
3.     return 0;
4. if y >= 1, then
5.     return (&k[x]+func(&k[x], &k[y-1]))
6. else
7.     return (&k[-x] + func(&k[x], &k[y+1]))
8. end if
9. End

```

## 3.2 Experimental Analysis

The experimental analysis has given us the idea about the functionality of the code and to practice evaluation accordingly. The above proposed method is smoothly evaluated on the 'Windows 10' operating system.

For evaluation, a sample code of multiplication is taken into consideration in the proposed technique, and majorly six to seven basic steps are required to build the program.

The experimental evaluation shows the perspective of the execution in the respective manner. The entire body of the code is encrypted in the C language, as we define the clear original code with none of the obfuscation applied and as the obfuscated code on which the same arithmetic operator functionality is applied (Table 2).

## 4 Limitations and Future Scope

The only disadvantage can be stated as that our proposed mechanism is limited up to specific code.

**Table 2** Statistics for correctness and time

Criterion	Multiplication	
	Original code	Obfuscated code
Execution Time	0.397	0.458
Output	4	4
Errors	No	No
Functionality Variation	No	No

Unfortunately, there is no perfect way to protect the software or the code from being cracked but by using necessary code obscuring methodologies to make an attack more time-consuming and to avert the attacker’s efforts.

Whereas the proposed mechanism proves to be functional and could be used as a subprogram of huge project. In a bigger picture, we look forward to make a spoonful of contribution by providing our multiplicative operator obscuring technique for future security of codes in general. Also, the obfuscation can be improved by using new patterns or features over different operators.

5 Conclusion

- To prevent the chances of insider attacks within the cloud and protect the ASCII text file by stealth, obfuscation on ASCII text file is applied. Obfuscating the whole ASCII text file is cumbersome through memory overhead and execution time.
- So, we are going to propose different data obfuscation techniques which concentrate on obfuscating the arithmetic operators (+, −, \*, /) on security essential part of ASCII text file.
- Here, we proposed the novel technique in data obfuscation by using reference pointer mechanism and successfully obfuscated the arithmetic operator multiplication (\*').
- There are many obfuscation techniques, and every one of them are often used permanently also as for bad purposes.
- Security teams encounter obfuscation a day once they either attempt to protect their code or de-obfuscate some malicious code originating from hackers.
- Achieving obscurity without creating any quite hindrance to the code’s actual functionality was the elemental motive behind all the minimizations and efforts were taken to form sure that the first code should not be suffering from any cause.

## References

1. Ahire P, Abraham J (2020) Mechanisms for source code obfuscation in C: novel techniques and implementation. IEEE, 2020. 978-1-7281-5263-9/20
2. Balakrishnan A, Schulze C (2005) Code obfuscation literature survey. In: CS701 construction of compilers, 2005
3. Tiella R, Ceccato M (2017) Automatic generation of opaque constants based on the KC lique problem for resilient data obfuscation. IEEE
4. Demissie BF, Ceccato M, Tiella R (2015) Assessment of data obfuscation with residue number coding. IEEE
5. Heffner K, Collberg C (2007) The obfuscation executive. In: Information security. Springer, pp 428–440. [5] Jakubowski MH, Saw CW, Venkatesan R (2007) Iterated transformations and quantitative metrics for software protection. In: SECURE, pp. 359–368. [6] Anckaert B, Madou M, Sutter BD, Bus BD, Bosschere KD, Preneel B (2007) Program obfuscation: a quantitative approach. In: QoP'07: proceedings of the 2007 ACM workshop on quality of protection. ACM, New York, NY, USA, pp 15–20
6. Xu H, Zhou Y, Kang Y, Iy MR (2017) On secure and usable program obfuscation: a survey. Cornell University Library, [arXiv:1710.01139v1](https://arxiv.org/pdf/1710.01139v1), <https://arxiv.org/pdf/1710.01139.pdf>
7. Yasin A, Nasra I (2016) Dynamic multi levels java code obfuscation technique. Int J Comput Sci Sec 10:140–160
8. Ahire P, Abraham J (2018) Perceive core logical blocks of a C program automatically for source code transformations. In: Proceedings of 18th international conference on intelligent systems design and applications. Springer, 2018. 978-3-030-16657-1\_36
9. Collberg CT, Low D (1998) Manufacturing cheap, resilient, and stealthy opaque constructs. In: Proceedings of the 25th ACM SIGPLAN-SIGACT symposium on principles of programming languages
10. Schrittwieser S, Katzenbeisser S Code obfuscation against static and dynamic reverse engineering. Springer J
11. Protect your C/C++ Code: Stunnix Obfuscator. <http://stunnix.com/prod/cxxo/>
12. OBFUSCATOR: obfuscate C & C++. <https://picheta.me/obfuscator>
13. The Tigress C Diversifier/Obfuscator. <http://tigress.cs.arizona.edu/>
14. Mangle-It C/C++ Obfuscator. [https://cnet.com/Mangle-It-C-Code-Obfuscator/3000-2229\\_4-190341.html](https://cnet.com/Mangle-It-C-Code-Obfuscator/3000-2229_4-190341.html)
15. Source Code Obfuscator: Semantic Design. <http://www.semdesigns.com/Obfuscators>
16. Obfuscator for Assembler Source Code. <https://www.pelock.com/obfuscator/>
17. CodeMorph C/C++ Obfuscator. <http://www.sourceformat.com/code-obfuscator-cpp.htm>
18. Zarate C, Garfinkel SL, Heffernan A, Gorak K, Horras S (2014) A survey of XOR as a digital obfuscation technique in a corpus of real data. DTIC Document, Technical Report, 2014
19. Demissie BF, Ceccato M, Tiella R (2015) Assessment of data obfuscation with residue number coding. In: IEEE/ACM international workshop on software protection

# Design and Implementation of Laptop Tracking System Based on Cloud Computing and IoT



Soumyadip Chatterjee

**Abstract** Here, I report the synthesis of laptop tracking and alerting system using AWS cloud serverless architecture and IoT keeping in mind the factors like laptop theft or performing unethical activities in laptop. Unlike existing laptop tracking techniques, the methodology mentioned in the paper is totally automatic, enhanced with new cloud technologies, maintains strong encrypted data transmission, and its performance is more optimized for the real need of organization facing problems to protect their systems and their cyber-ethics especially in the WFH schedule. Even though the organization follows a strong protocol in monitoring their systems usage inside their office network but when people started working from home, they are no more connected to office network which leads to many blind spots to maintain their monitoring protocols. This paper explains the implementation of fetching large chunk of crucial data from any laptop irrespective of their operating system. Using just a click away the design is capable of setting up all prerequisites in the host laptop and after transmitting all data it will clean up all its presence from the host laptop. All operations takes place in few minutes.

**Keywords** Laptop tracker · Cloud computing · Internet of Things · Serverless architecture

## 1 Introduction

In simple terms, we can define laptop as a computer which is portable and suitable to use especially while travelling. In IT industries, sometimes employees get a desktop computer with a predefined desk or a laptop which makes employees to complete their tasks in the dwelling place. Unlike office properties like desktop computer, laptops are hard to keep secure and safe as its mobile and easily concealable. Organizations transitioned their employees from the office to a work from home model due to the COVID-19 global pandemic. Almost every organization, they try to follow a strict

---

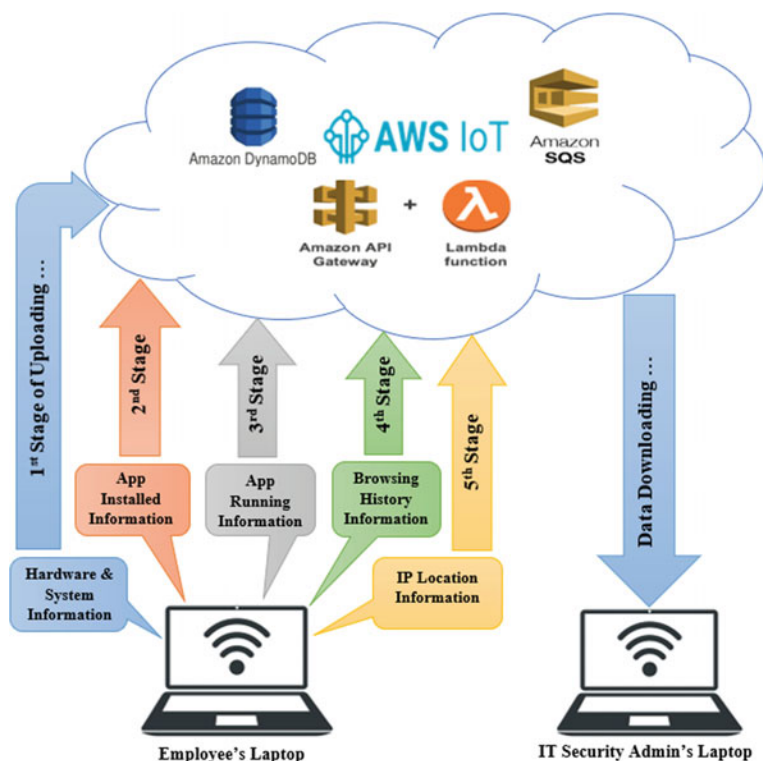
S. Chatterjee (✉)

Edifixio India Private Limited, AF Block (Newtown), Action Area 1A, Newtown, Kolkata 700156, West Bengal, India

IT policy inside the office premises to prevent their system getting used for unethical purposes. When work from home model got implemented, employees started using their office laptop at home which made them vulnerable to the issues like laptop theft or performing unethical tasks as they are using their own private network. IT security admin team of some of the organizations finds it difficult to execute their monitoring protocol in order to follow their IT policies. Rather than scheduling any fix disclosable dates for auditing the laptops, it is better to use an automated fast system which can not only scan every laptop at same time without letting the employees know about getting monitored but also can be executed anytime and as many times the IT security admin team wants. It will be also beneficial in not interrupting any employees work flow. In such scenarios, the laptop tracking system is developed by exploring multiple applications to prevent the issues like laptop theft or employees of any organization to perform malicious activities inside their system especially during WFH schedule. This is an automated data logging system which is developed using enhanced technologies which is followed by a strong data encryption policy. The whole operation of data logging system takes to complete approximately less than 5 min, beginning with setting up all prerequisites and ending up cleaning up all its traces inside the host laptop after sending all data to its target. All data are stored using AWS cloud services in a segregated and secure way. Finally, IT security admin team or the monitoring team can do all the analysis and presentation of the recorded information in a graphical way, and they can also save the user data from cloud to their local system in multiple formats. The analysis can be done either using AWS GUI mode or CLI mode as per user choice [1–3].

## 2 Proposed System

The system is mainly the communication between two laptops, and the medium of operation takes place through cloud computing and IoT. This application is designed in a serverless architecture model where an automated script is executed into a host laptop/employee's laptop. The automated script sets up all the prerequisites in its first go based on its operating system. As soon as script is ready with necessary modules, it does initiate another script to proceed with which is responsible in transferring the data. All the important details referring to that laptop are carried away in five different stages with an unnoticeable gap into the AWS cloud. In the overall operation, there is a large chunk of data that are transported into the cloud databases and into the IoT shadow. Every time the script is executed, a unique user is created with the following unique id structure *"logged in username\_desktop name\_current date"*. If the script is executed on day basis, it will create a new user with unique id, but if it is executed for multiple times in the same day, then the existing user data will get overwrite with the current data. The application segregates all data into five following stages while transferring them (Fig. 1):



**Fig. 1** Basic block diagram of the laptop tracking system using cloud computing and IoT

1. **Stage 1: Hardware and System Information**—During this stage, all the hardware-related information is transferred. In this segregation, all the data are sent in JSON structure. In this stage, a laptop's system information, boot time information, CPU information, memory information, disk information, network information and GPU information are all transported into AWS Dynamo DB.
2. **Stage 2: App Installed Information**—During this stage, all the details of currently installed applications are transferred into AWS Dynamo DB which include its app name, version name, install date, install location and publisher name. In this segregation, all the data are sent in JSON structure.
3. **Stage 3: App Running Information**—During this stage, the current running application's details are transferred into the AWS Dynamo DB which include its name and PID number. In this segregation, all the data are sent in JSON structure.
4. **Stage 4: Browsing History Information**—During this stage, latest one thousand browsing histories are transferred into AWS Dynamo DB which include the URL, Web site name and date–time of visit. The tracking system is capable of extracting data from the browser like Chrome, Mozilla, Safari, Opera, etc.,

and it is an OS independent process. The range of browsing history can be altered up and down. In this segregation, all the data are sent in JSON structure.

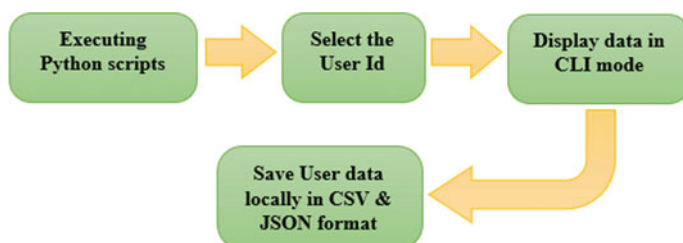
5. **Stage 5: IP Location Information**—During this stage, laptop location is traced, and its data are reflected into the AWS IoT Core shadow. The data include laptop name, user public IP address (private IP already got transferred in stage one under network information), connected broadband name, current living city, state, country, postal code latitude, longitude and time zone. So, if the location is not matched with the desired one, then security IT team immediately gets an alerting email with all its details.

As it is mentioned earlier, the whole application is running by multiple automated script, and the process is actually initiated by a batch script. After completing sharing all the data of the host laptop, all the traces of its execution and all the executed file are removed from the host laptop using its auto-deletion mode. All operation generally takes place in less than five minutes (see Fig. 3).

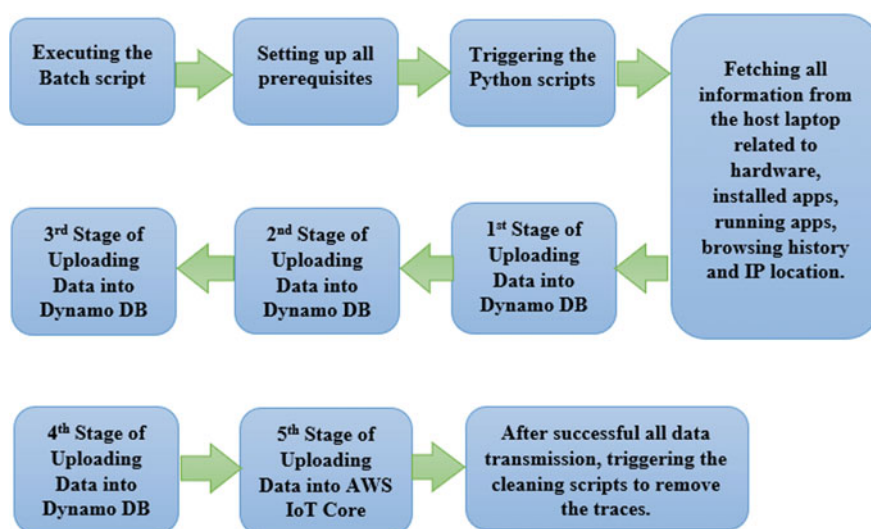
Every stage follows a strong data encryption policy while sending data. In every stage, all the data are converted into an encrypted token which are sent to the Amazon API Gateway. Every time API Gateway receives a request, it verifies every token using custom authorizer in three steps. The steps include decoding the token using secret key and special algorithm, verifying the headers of token and signature validation. If any steps fail to verify which indicates a middle man attack that specific request is get blocked by the custom authorizer and won't let access the lambda code. If the request is valid, API Gateway will trigger the AWS Lambda which is a serverless compute service that lets run code without provisioning or managing servers. AWS Lambda is computing every request and based on the request lambda decides its endpoint. For example, for first four stages of data sharing, lambda is responsible for computing the data and storing them into Dynamo DB, whereas in the last stage of data sharing where location information is transferred, lambda performs all necessary computation followed by it storing those data in AWS IoT Core shadow. AWS IoT Core is uniquely integrated with Amazon SQS service using IoT Rules Engine. IoT Rules Engine is responsible for checking the reported data with the desired data in the IoT core shadow, if anything is not match, it will immediately trigger the SQS service which will send an email to the admin team. In this case, Rules Engine is checking the location information from a prefixed location information, so if the host laptop is detected outside the desired location/city, it sends an email to the security admin team with the details of the laptop (see Fig. 4) [4].

After data collection, security admin team can check their systems data in two ways which include GUI mode and CLI mode. For GUI mode, they can access the data using AWS console. For CLI mode, there is a script to extract data from cloud. Security admin team needs to execute the script in their system and followed by that they can see all data in their console after selecting user id. Auto-download feature is auto-enabled, and therefore, all the data which the security team is currently watching on screen are already downloaded in JSON & CSV format inside a unique name folder based on its user id (see Fig. 2) [5].





**Fig. 2** Flowchart of data display in CLI mode



**Fig. 3** Flowchart of sending laptop data from host laptop into AWS Cloud

## 2.1 Flowchart of the Proposed System

## 3 Results and Discussions

Above images show the databases which are the final endpoint for first four stages of sending data. There are all total five NoSQL databases, among them, four are dedicated for storing the data which I have mentioned above; but one database is responsible for storing all IoT connectivity details of the user along with its creation time (see Figs. 5, 6, 7, 8 and 9). The data include certificate ARN, certificate id, certificate string private key, etc. If security admin team wants to access the data using GUI mode, then in this way through AWS Dynamo DB console, it is possible.

If network admin wants to access the location details using GUI mode, then they can access it using AWS IoT Core console (see Fig. 10). If the location is not within

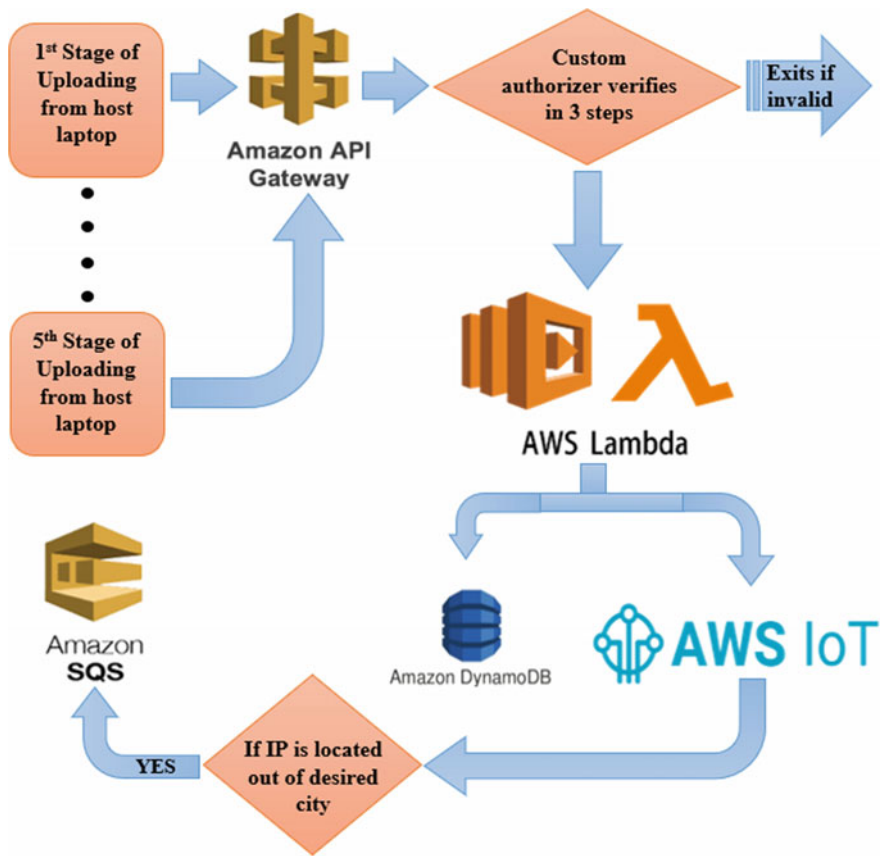


Fig. 4 Flowchart of data computation in AWS Cloud

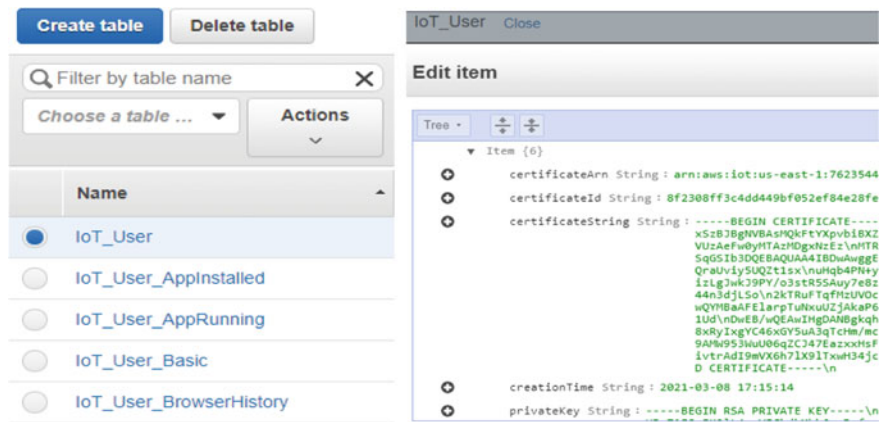
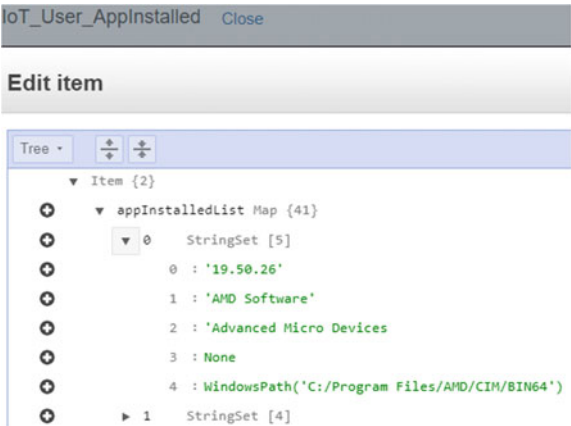
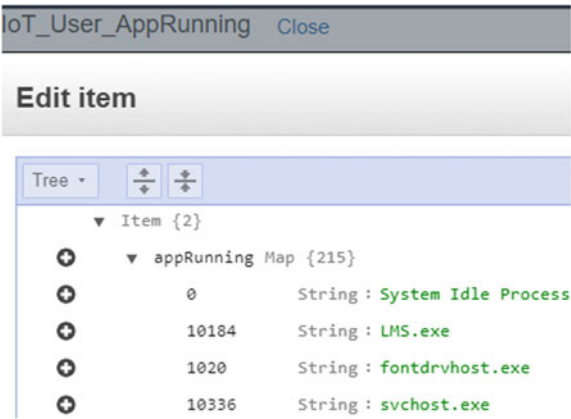


Fig. 5 All the tables (left), Dynamo DB table for IoT security details (right)

**Fig. 6** Dynamo DB table for Stage 2



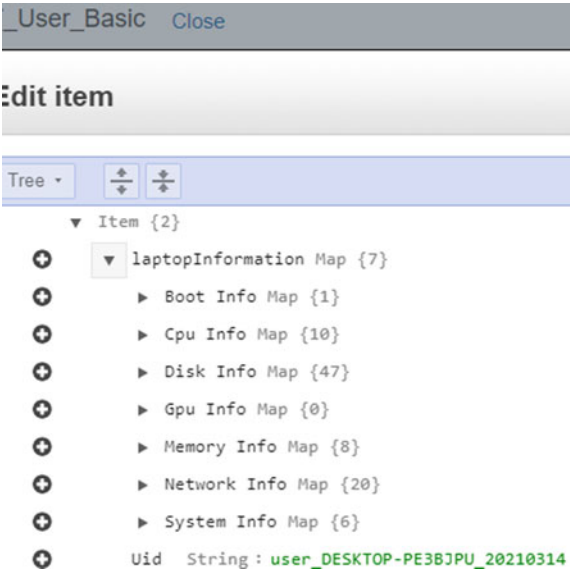
**Fig. 7** Dynamo DB table for Stage 3



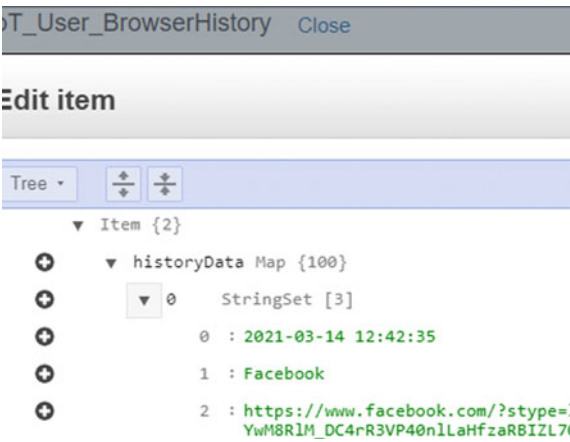
the set perimeter, then eventually, they will receive an email with the IP address and its corresponding location details (see Fig. 11).

Above images show the output of data accessing using CLI mode. From the dropdown, user id needs to be got selected, and once it is selected, the automated script starts downloading all the data of the specified user id from IoT Core and Dynamo DB (see Fig. 12). The data get saved inside an auto-created folder renamed as the user id, and the files are renamed with respect to databases name. The downloaded files are available in JSON and CSV format both (see Fig. 13). After the download, all the data are extracted and displayed on screen in table format using CLI mode (see Fig. 14). CLI mode stands for command line interface, and GUI mode stands for graphical user interface. Both the modes are provisioned to make it more user-friendly. CLI mode to be used for the user using console in Linux/Windows/Mac and GUI mode to be used for the user using graphics to interact with an operating system.

**Fig. 8** Dynamo DB table for Stage 1



**Fig. 9** Dynamo DB table for Stage 4

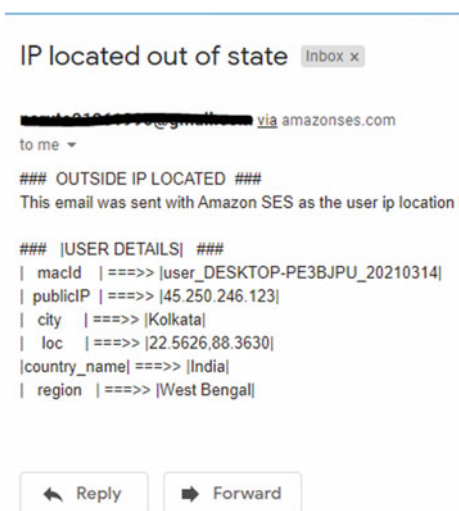


4 Conclusion

The application is actually built to protect the cyber-ethics and the corporate IT policies. I have maintained the cyber-ethics throughout the application in a proper and safer way. In today’s WFH implementation, any organization’s infrastructure admin team is facing problems in tracking not only employees’ system usage but also performing auditing without interruption or delaying any employee’s project delivery time. In order to audit all employees’ system in a speedy process, they

**Shadow state:**

```
{
  "desired": {
    "expected_city": "Bangalore"
  },
  "reported": {
    "macId": "user_DESKTOP-PE3BJPU_20210314",
    "publicIP": "45.250.246.123",
    "hostname": "node-45-250-246-123.alliancebroadband.in",
    "city": "Kolkata",
    "region": "West Bengal",
    "country": "IN",
    "loc": "22.5626,88.3630",
    "org": "AS23860 Alliance Broadband Services Pvt. Ltd.",
    "postal": "700006",
    "timezone": "Asia/Kolkata",
    "country_name": "India",
    "latitude": "22.5626",
    "longitude": "88.3630"
  }
},
```

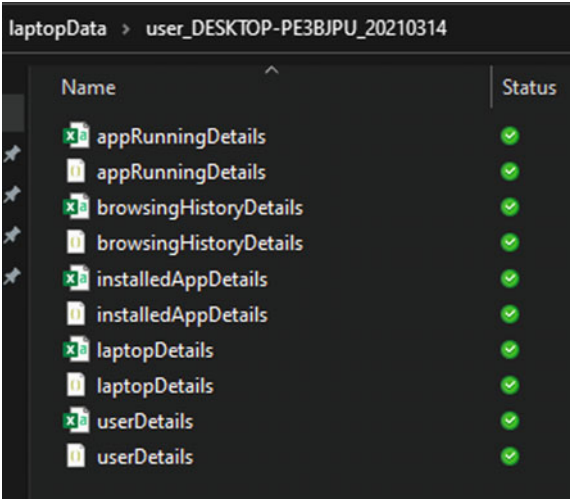
**Fig. 10** IoT shadow consists of details of Stage 5**Fig. 11** Email sent if IP is out of state

need to be in the office network or share common network at least, and if any automated process is not followed, then employees might have to wait physically for the auditing to finish, which might lead to work delivery delay. This application can be executed anywhere irrespective of their employees' using shared network. Not only this application is almost OS independent but also there is nothing browser-specific while data fetching. The application can be enhanced into such a level that it can keep monitoring your keypad events, accessing the web cam to get the biometric details,

```
? = Choose user's id to see details = user_DESKTOP-PE3BJPU_20210314
Directory named 'user_DESKTOP-PE3BJPU_20210314' already exists
export dynamodb: IoT_User_Basic
Connecting to AWS DynamoDb
Downloading 2 records ..... 100%
Total downloaded records: 2
Writing to csv file.
export dynamodb: IoT_User
Connecting to AWS DynamoDb
Downloading 2 records ..... 100%
Total downloaded records: 2
Writing to csv file.
export dynamodb: IoT_User_AppInstalled
Connecting to AWS DynamoDb
Downloading 2 records ..... 100%
```

Fig. 12 User id selection and downloading its data

Fig. 13 Downloaded files



extracting of username and passwords, etc., but following cyber-ethics, nothing as such is implemented.

The application is compact, easy to use, follows strong data encryption policy, efficient and most importantly runs as serverless architecture which is totally managed by AWS. So, there are almost negligible chances of any downtime, and billing prices are also cheaper compared to other laptop tracking implementation, especially while using any other cloud services.

For future enhancement, I am thinking for few changes listed below:

- i. Decreasing the number of clicks to execute this automated process by implementing the auto-triggering mechanism like clicking on the auto-download link will automatically trigger the script execution.

IP LOCATION PARAMETERS		VALUE
Expected City	=====>	Bangalore
Mac ID	=====>	user_DESKTOP-PE3BJPU_20210314
Public IP	=====>	45.250.246.123
Hostname	=====>	node-45-250-246-123.alliancebroadband.in
City	=====>	Kolkata
Region	=====>	West Bengal
Country Code	=====>	IN
Location	=====>	22.5626,88.3630
Internet Service Provider	=>	AS23860 Alliance Broadband Services Pvt. Ltd.
Timezone	=====>	Asia/Kolkata
ISP Postal Address	=====>	700006
Latitude	=====>	22.5626
Longitude	=====>	88.3630
Country Name	=====>	India

LAPTOP PARAMETER	VALUES 1	VALUES 2
Mac ID	user_DESKTOP-PE3BJPU_20210314	
Disk Info	Used 2	5.56GB

Fig. 14 Display output in CLI mode

- ii. All the content of the files should be encrypted so that if the files are shared unethically, it will restrict itself from getting exposed.
- iii. Implementing employees’ biometric authorization using a web cam.
- iv. Listening to employees’ key press action, monitor or store any unprofessional word usage with timestamp. It will also help in solving any escalation regarding abusive language usage.

References

1. Datta N, Malik A, Agarwal M, Jhunjunwala A (2019) Real time tracking and alert system for laptop through implementation of GPS, GSM, motion sensor and cloud services for antitheft purposes. In: 2019 4th international conference on internet of things: smart innovation and usages (IoT-SIU), pp 1–6

2. Pushpa KG, Sharanappa K, Pramod SN, Prashanth CH, Sujithkumar C (2020) Laptop tracking and alert system using GPS and GSM module. Int Res J Eng Technol (IRJET) 7(6):3012–3015

3. Narayana TVS, Srikanth SA (2013) Laptop tracking mechanism using GSM/GPS technology. Elixir Int J 60:16214–16218

4. Rashik A, Simi KS, Esahak IM, Razak MNA (2017) A predictive controller for mobile/laptop tracking and security. Int J Innov Res Sci Eng Technol 6(5):74–78

5. Mhlaba A, Masinde M (2015) A hardware based model for an asset monitoring and tracking system: case of laptops. IEEE, pp 1–7

# Design and Implementation of Women Security System Based on GSM and GPS Technology



Soumyadip Chatterjee and Surajit Mandal

**Abstract** Here I report the synthesis of Women Security System using GSM and GPS technology. During this communication, a sensible security system is projected so as to boost the safety of women, minors, and less secured persons. So, for their security and safety purpose, government has provided security through rules and regulation to the society. Although there are many existing systems for security purpose, they lack of advanced smart security system, so as to beat such issues advance security system for women is enforced. This project ensures about the safety for women whose design and implementation comprises of an Arduino controller and sensor. The main purpose of our project is to provide security to the women from the harmful places. During this project, we tend to AR providing facility to secure the women by providing this kit. A vibrator motor, accelerometer, GSM, and GPS are utilized in this project. Once the women are in threat, the device senses the movement through accelerometer. Once the sensor crosses the brink limit, the device gets activated and traces the location of the victim using the GPS module. By using the GSM module, the victim's location is sent to the predefined contact number which is already saved in program.

**Keywords** Falling-off detection • Location finder • SOS

## 1 Introduction

Women safety in India has become a major concern since last few decades. The line graph of crime rates against women is incremental. The worst fear of their safety is the sad reality of our country. Today, women empowerment is on rise but the hideous crime rate against them makes it conceal. Our country name is a feminine given name where women are given stature of goddesses but still you will find details of crime committed against women every day in your newspaper. Technological advancement

---

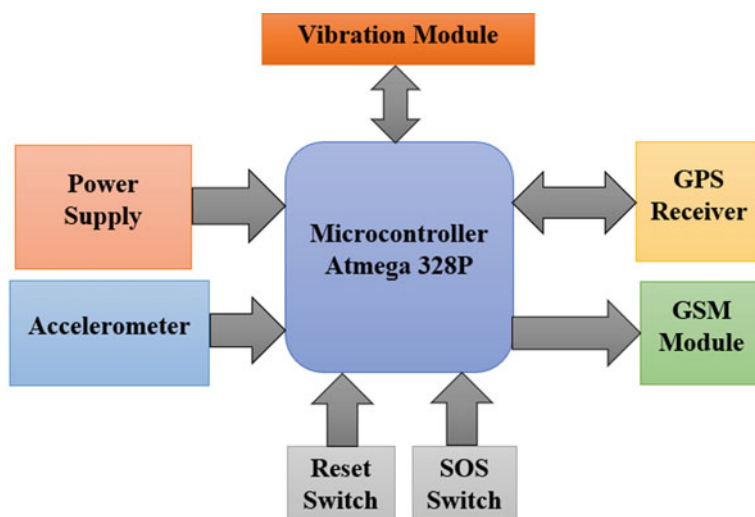
S. Chatterjee (✉)

Edifixio India Private Limited, Kolkata, West Bengal 700156, India

S. Mandal

B.P. Poddar Institute of Management and Technology, Kolkata, West Bengal 700052, India





**Fig. 1** Systematic block diagram of the system

has always been a matter of pride for our country when it comes to progress. Today, there are security and surveillance system for every possible domain required but we have not seen much technological advancement when the matter is related to woman's security concern. So, we have gone through few previous researches on this topic and came into idea of designing a system with minimal loopholes and minimizing the false alarming rate ratio. We have tried to implement an enhanced idea of designing a customized, accurate, robust, and a compact system using GSM-GPS technology for women's safety [1].

Even during epoch women are feeling insecure to step out of their house because of increasing crimes in our country like harassment, abuse, violence, etc. Many women are operating in company even in night shifts. There is a sense of insecurity among the female employees. The projected device is lot of sort of a safety system just in case of emergency. This device can be fitted in a very jacket or blazer for girls. It is straightforward to hold device with more features and functions. The emergency push button is held to one of the buttons of the blazer. The main purpose of this device is to intimate the parents and police about the current location of the victim. A GPS system is utilized to trace the victimized spot of the victim, and a GSM module is employed to send the message to the predefined numbers. This model is additionally helpful for little children and senior aged individuals conjointly [2, 3].

## 2 Proposed System

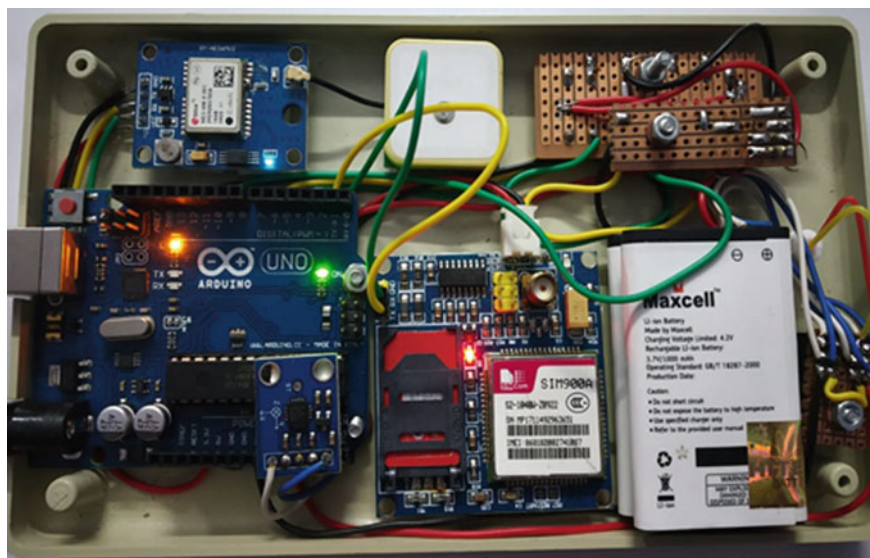
As shown in Fig. 1, Arduino is used as main microcontroller, and this system is made for incident alert; the whole system is to be fitted in a jacket itself (similar to a blazer for women). So, once the accident happens, the vibration module is triggered to assure user its activeness. With GPS the latitude and longitude of that particular location are obtained which helps in discovering the precise location of the accident site and followed by GSM modem SIM900A which is interfaced with microcontroller. So that, when accident happens, the GSM module will be sending SMS automatically to the particular numbers which would be entered in the code. Accelerometer will check the shock intensity and the validity of sending message. Also, the delay time is programmatically integrated to give the person chance to press the reset key if the threat detection is normal which will reduce the false alarm ratio. SOS key is also interfaced with the microcontroller to send alert manually.

In the frontal side (see Fig. 2), the system is having three switches and a charging point. SOS and Cancel button have been set closely intentionally. After pressing the SOS button for 2 s, it will trigger the alert and within 10 s message will be sent but if the user pressed it by mistake or maybe she misjudged the situation, in order to save from false alert, the cancel button has been set closely to SOS button (Fig. 3).

This project ensures concerning the protection for women which contains of an Arduino controller and sensor. A vibrator motor, accelerometer, GSM, and GPS are used in this project. When the women are in threat, the device senses the movement through accelerometer. When the sensor crosses the brink limit, the device gets activated and traces the situation of the victim utilizing the GPS module inside the system. By using the GSM module, the device traces out the victim's location and an alert is sent to the registered contact number with the location coordinates; and it is mandatory to insert a fully active SIM card in the GSM module.



Fig. 2 Frontal view of the system



**Fig. 3** Internal circuit of the system

There will be two buttons: one is for sending SOS to the configured number and one is for canceling it. Two switches will be too small but finger reachable. Whole system is as small as a woman's purse. After pressing the SOS switch for approx. 2 s, an emergency alert "Save Me" message will be delivered to the configured receiver within 20 s. There is a vibration module, which will keep on vibrating for 20 s while sending SMS so that user remains aware of its system functionality. SOS message will be sent along with its longitude and latitude. Link will be there in SMS which will directly open into Google Map. Now, an accelerometer is being implemented so that any type abnormal gesture (fall over, back push, etc.) can detect and respond automatically within that same delay. With this feature, if a woman is attacked from behind and loose her grip from her purse, then this system will be useful. So, pressing SOS button is not mandatory, our designed system is semi-automatic. This system is not only portable but also rechargeable. One-time recharge and the system will be functional for next 4–6 h. External power supply is also present.

Here in Fig. 4, the circuit connections of this project are displayed. The Tx pin of GPS module is connected to digital pin (D10) of Arduino Uno. We have used Software Serial Library to allow the serial communication on pin D10 and D11 and made them Rx and Tx, respectively. Generally, Pin 0 and 1 of Arduino are utilized for serial communication but by using the Software Serial Library, we can accommodate any digital pins for serial communication of the Arduino. 3.3 V supply is used to power the GPS module [4].

GSM module's Tx and Rx pins of are directly connected to pin D0 and D1 of Arduino for GSM interfacing. GSM module is also powered by 8v supply. An accelerometer is added in this system for detecting an incident, and its x-axis, y-axis,

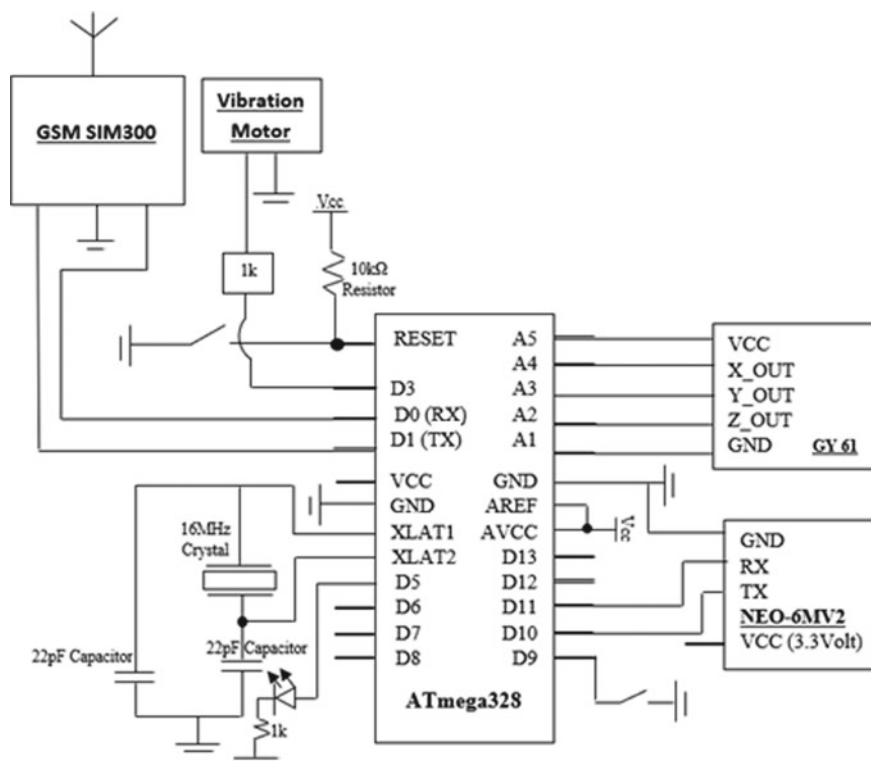
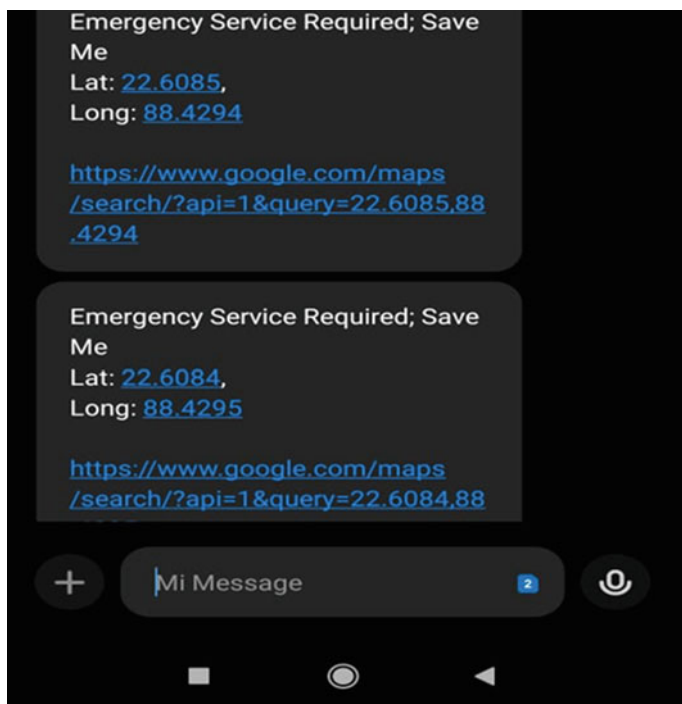


Fig. 4 Circuit diagram of the women security system

and z-axis ADC output pins are directly connected to Arduino ADC pin A1, A2, and A3. A vibration motor is attached in this system for assuring the carrier that is the device is working all fine and its wires are connected to D3 and GND pin of Arduino, respectively. Key work as a switch for reset the circuit when there is no need for the ambulance or the police [5].

### 3 Results and Discussions

The ultimate output of the system will be in the mobile as per settings made by the user. The preset setting is the emergency contact to set where the SMS will be sent. In the SMS, a customized message will be there which presently state “Emergency Service Required; Save Me” followed by the latitude–longitude of the victim spot with a link which is integrated with Google Map (see Fig. 5). So, when the person gets the SMS, he/she can know the location of the victim in an instant click.



**Fig. 5** Output of the system in the mobile phone

## 4 Conclusion

The proposed system will handle critical issues related to women's safety in the future and will help to solve them with technically advanced equipment and concepts. This enhanced design will defeat the fear that scares every woman in the society regarding her safety and security. Our motive is to design and build a gadget which is so compact and portable that it will be advantageous to personal security system or the emergency response system which will be useful for women in the incidents of crime. The proposed system is made in a cost effective and efficient which can store the information of the members in the particular locality and provide immediate alert in case of crime against women.

For future enhancement, we are thinking for few changes listed below:

- i. Increasing the storage ability to hold a greater number of contacts to send alert during the emergency time.
- ii. Corresponding to the SMS, place a demo call to the receiver to alert him more profoundly.
- iii. Upgrading to biometrically triggered or voice activated system.
- iv .Upgrading to next level of miniature system.

## References

1. Seelam K, Prasanti K (2018) A novel approach to provide protection for women by using smart security device. In: 2nd international conference on inventive systems and control (ICISC)
2. Dhole SV, Ramraje SV, Bhosale BB, Jagtap G (2019) IOT based accident tracking and reporting. *Int Res J Eng Technol (IRJET)*, 6(12)
3. Nadim M, Aravindh S, Qayyum Khan R (2018) An automated fall detection system using accelerometer. *Int J Sci Eng Res (IJSER)* 9(7)
4. Ms.Sonali S. Kumbhar, Ms.Sonal K.Jadhav, Ms. Prajakta A.Nalawade,Ms. Tamanna Y.Mutawalli: WOMEN SECURITY SYSTEM USING GSM AND GPS: *International Research Journal of Engineering and Technology (IRJET)*, volume 5, issue 3 (2018).
5. Evangelin C, Christina M, Malapriyadharshini I, Priyadharshini G, Kowsalya S (2019) Women safety system device using GSM module. *Int J Emerg Technol Innov Eng (IJETIE)* 5(9)

# A Scalable Framework for High-Performance Computing with Cloud



Manish Kumar Abhishek and D. Rajeswara Rao

**Abstract** In cloud environment, virtual machine (VM) is provisioned depending on the configuration provided for the hardware computing resources based on user requirement. In order to provision these VMs, cloud environment has its own distributed infrastructure of several servers. On major and minor scale, infrastructure as a service is getting used as a feasible platform for the execution of high-performance computing (HPC) applications. However, few lags still exist within the futuristic traits of HPC application and current cloud scheduling-based algorithms. This paper presents a scalable framework that uses an algorithm-based approach to predict the matching underlying computing resource suitable to run high-performance computing (HPC) applications. It will behave like a decision-maker to predict the placement of containers. It is developed to facilitate a cloud computing environment to provide solution for availing the appropriate computing resources from dedicated allocated containers for HPC and monitoring of the same using Kubernetes. The framework along with proposed solution over containers is discussed and illustrated by computational algorithm with results. It will scale up or down the resources as required and will spawn or decommission the HPC clusters within minutes. We use the Kubernetes to monitor the deployed containers, and our results indicate that shared resources are having no impact on performance.

**Keywords** Cloud computing · Container · Docker · High-performance computing · Kubernetes · Virtual machine (VM) · Job scheduler

## 1 Introduction

The key concept of cloud computing relies on virtualization. The cloud servers are the virtual servers running in cloud environment which can be built, hosted, and accessed remotely using cloud computing platform via Internet that can be provided as a part of infrastructure, application, or software service. Virtualization addresses resources

---

M. K. Abhishek (✉) · D. Rajeswara Rao  
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India  
e-mail: [manish.abhishek@gov.in](mailto:manish.abhishek@gov.in)

via leveraging encapsulation, isolation, and partitioning. It helps in cloud computing for provisioning computing instances on demand as per user requirement in form of virtual machine (VM). Virtual machine runs on hypervisor and consumes resources in terms of memory, CPU, network bandwidth, persistency from physical hosts that are the part of cloud infrastructure. Every VM needs a unique guest operating system (OS) having its own binaries, libraries where multiple VMs can run on one physical server with different OS where, on the other hand, containers gain a popularity via providing OS virtualization so that multiple workloads can easily run on a single instance of OS and results into agility, speed, and portability. In summary, containers are very light weight and easily portable.

High-performance computing (HPC) applications are widely used in research, laboratories, and in business of analytics and cloud computing will act as a cost-effectivity and workload sharing solution since users do not need to worry about cluster creation, start up, and maintenance. In addition to that, cloud will offer elasticity and scalability in terms of computing resources. It provides benefits of visualization to HPC applications. Earlier, clouds have been proposed for deploying and hosting Web applications, where computing resource specifications were not in line with HPC applications. HPC applications require high bandwidth along with low latency interprocessor communication for gaining better performance. In cloud, somewhere it results into a bottleneck for HPC applications. Supercomputers and mainframes have input–output subsystems and operating system (OS) for matching the criteria of HPC applications. However, some efforts in cloud computing area for HPC, such as Nebula [1] and Magellan [2], look like promising from research and analytics perspective. Earlier, research [3–5] in HPC applications execution on cloud has been dedicated to achieve performance in terms of metric, and studies were focused majorly with Amazon EC2 [6] on miniscale. The results were not satisfactory, and even after having the huge advantages of cloud computing, it cannot be used for all HPC application execution.

In this paper, we have presented a scalable framework that uses an algorithm-based approach to predict the placement of matching underlying computing resource suitable to run high-performance computing (HPC) applications on cloud environment. Containers replacing the physical machine will be a significant impact on performance, and the primary issue that we address through this research is the improvement in HPC application performance on cloud via container placement strategy. Existing cloud lifecycle management systems such as Open Stack and Eucalyptus cannot be recommended as a scheduler to run HPC applications. We are proposing scheduling as a selection of servers for provisioning containers by utilizing a shared pool of computing resources. Here, we analyze the challenges and explore the alternatives for scheduling HPC applications on cloud. We focused on the implementation of HPC-based container placement approach computing resources awareness in algorithm-based scheduler and evaluate their performance measurement on cloud stack. This approach will not benefit only HPC but also the cloud users. Using this, cloud vendors can obtain higher profits for their computing resources. The main contribution of our proposed approach is having a novel scheduler for the scheduling and execution of HPC applications in cloud environment using containers via Docker



and orchestration through Kubernetes. The rest sections of this paper have been categorized in following manner: Sect. 2 describes the containerization background, Docker, and Kubernetes. Section 3 describes our proposed algorithms along with implementation. Section 4 is all about having the evaluation methodology. Section 5 is going to have performance results. Concluding remarks are provided in Sect. 6.

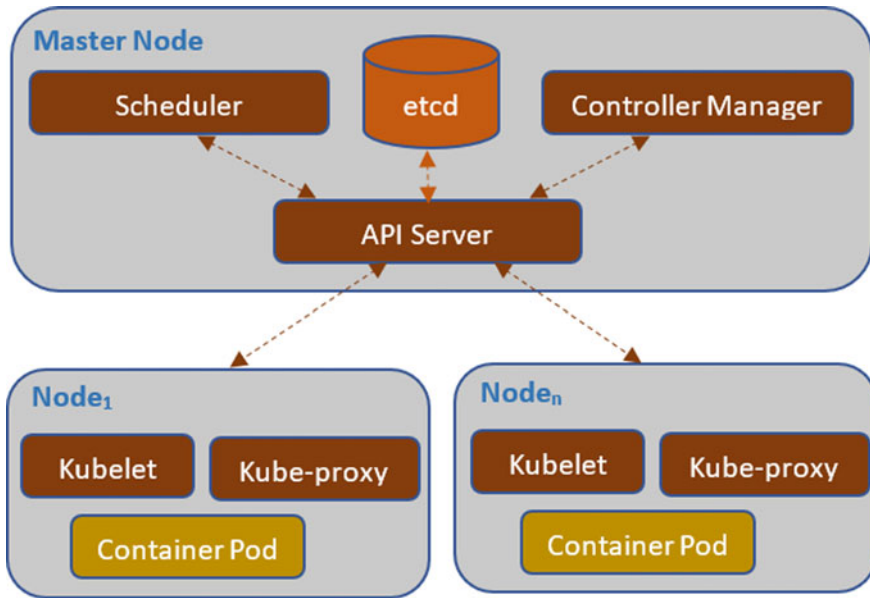
## 2 Containers, Docker, and Kubernetes

Containerization is an alternative of virtualization. It creates, packages, distributes, deploys, and executes application in a lightweight standardized execution environment called as container. Container runs on physical server via sharing the host OS kernel, its binaries, and libraries. It helps on run isolated systems on single OS host and reduces management overhead. They are very lightweight and more portable than virtual machines. This is one of the main reasons containers are gaining more popularity in world of virtualization. Docker is mainly composed for the lightweight containers having operating system virtualization as a core component in underlying layer for implementation, deployment and shipping of bundled applications having inbuilt dependent tools and libraries for application management. The architecture is based on client–server type including three major components i.e. Daemon process, CLI client and REST API. Commands need to be executed from the CLI client to communicate with the server via interacting with the daemon process running on the server. REST API can be used to communicate with daemon and to get the needful computing resources. It is responsible for building, running, and distributing containers having tagged images using service registry. It can be executed either on CLI Client on daemon server or on remote server (Fig. 1).

Kubernetes is used for automating the deployment, scaling, monitoring, and management of containers in terms of pods. In short it is container orchestration. It categorizes the containers to make application in logical unit for a better management and prediction of discovery. Here, we have used our own Kubernetes cluster for our containerized stack. It is loosely coupled and manages the workloads efficiently. It will unify the cluster of hosts in single pool of computing resources. Kubernetes will organize the running application in group of containers using Docker engine and manage it as per defined configuration. It will compute scheduling, self-heal, load balance, rollback, roll out, and scale service horizontally.

## 3 Container Scheduler for HPC Application

In this paper, we have proposed scalable framework having the container placement approach. Using this, we are addressing the problem of container placement to run HPC application, i.e., matching the criteria of n number of computing instance resource requirements like CPU, GPU, memory, and disk in terms of containers.



**Fig. 1** Kubernetes cluster using Docker containers

Map  $i$  containers ( $c_1, c_2, \dots, c_i$ ) to  $M$  number of physical servers  $S_1, S_2, \dots, S_m$  are already partially occupied or unoccupied at all. Moreover, our goal is on facilitate the cloud providers/users a container placement for HPC. In next section, we will cover the proposed design and its implementation.

### 3.1 Proposed Approach

HPC application is going to have parallel processes in count of  $n$ , which will mainly perform the interprocess interaction for overall execution. It is mainly based on cluster of servers that is quite unknown in cloud environment. The aim to cluster the containers is suitable to run HPC application based on resource requirement. Let us assume that in a cluster, each server is 8 core nodes, and there are 64 nodes in the hosted rack with a 20Gbps connection link to a switch. All hosted rack servers are interconnected with 20 Gbps link with a switch. Here, 20 Gbps link is shared in between 64 nodes, efficiently giving bandwidth of  $20\text{Gbps}/64$  equal to 0.625 Gbps between the placed two nodes in that belongs to another rack where bandwidth among two nodes that are residing within a single rack is 1Gbps. Therefore, it is suitable to cluster the containers to nodes within a rack instead of picking any policy at random for container placement which can distribute the nodes across the racks cluster. HPC applications have generally a tendency of iterative and massive synchronous, which

implies that in every repetition—computation is going to be followed by synchronized communication. The next iteration which will start once first will get over which will result it into a degradation in performance of application execution. In cloud, user is unknown to the underlying infrastructure using which containers have been provisioned. Cloud environment consists of heterogeneous physical servers to host containers. Existing container placement approaches ignore the heterogeneous nature of underlying hardware that results in some containers running on slower processors and non-GPU-based resources, while some running on faster and appropriate ones. Some cloud providers, for example Amazon EC2 [1], addressed the issue of heterogeneity via opting the creation of new compute instance and allocation on the basis of that. With a newly created compute instance, it is going to utilize the rest of the capacity and will allocate it to a new computing instance using shares (e.g., 70–30 CPU share). However, for HPC application, this will downgrade the performance factor as all processes of an application may be run out of sync because of placement of another instance on this node. To be instances in sync, all needs to be scheduled in a single run if they are executing on multiple platforms and sharing computing resources like CPU with other containers (gang scheduling). Apart from this, the arising interference is going to be another factor which will impact performance. To overcome this kind of interference, Amazon EC2 is recommended using a dedicated cluster for high-performance computing applications [6] but it will result in higher price. For providing the homogeneous infrastructure for containers to run HPC application, we take an alternative proposed approach. We prepared the container placement-specific computing resources that are going to be aware with underlying hardware and make sure that all containers for a request will get allotted on respective processors type as configured (Fig. 2).

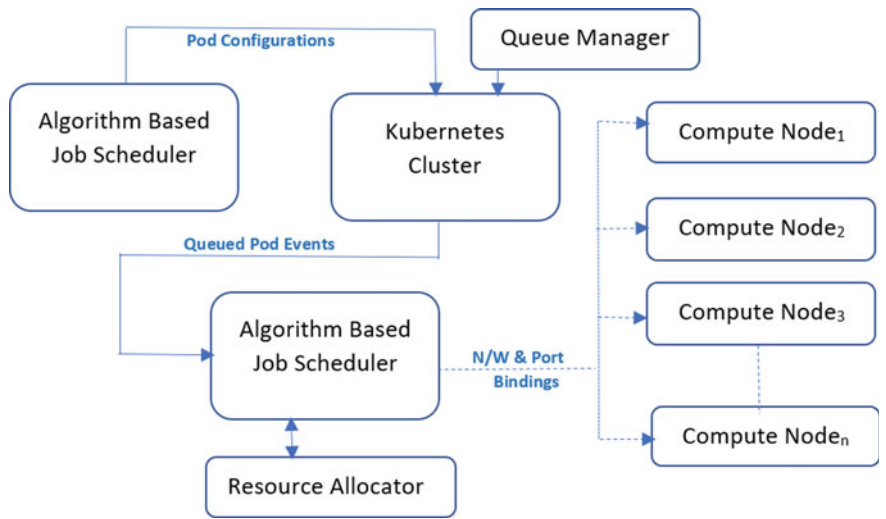


Fig. 2 High-level architecture based on proposed scalable framework

### **3.2 *Implementation***

We have implemented the proposed approach using our own stack of cloud to run HPC application. The first consideration is to place the (i) number of containers in the same cluster of rack in spite of its placement to distributed one. This consideration is defined as a part of Algorithm 1 (pseudocode) followed by having the list of physical servers. The second phase of refinement will consider only the GPU-based containers to run the HPC application and will remove the unmatched ones based on defined configuration. It will also compute the count of matching containers that can fit in same server (on the basis of CPU, memory, cores). We will call it serverQuota. Next, we will have the count of containers, and the rack can have (rackQuota) by adding the serverQuota of all servers in a rack.

**Algorithm 1:** pseudocode for the container scheduling**Input:**  $S_n, NH_n, H_n, R_s, SQ, HQ$ **Output:**  $PP_m$ 

```

1:  $S_n = \varnothing$ 
2:  $NH_n, H_n = \text{new vector} \langle \text{int} \rangle$ 
3:  $R_s = \text{new Set} \langle \text{int} \rangle$ 
4:  $SQ, HQ = \text{new int} \langle S_n \rangle$ 
5:  $\text{serverCapacity} = \text{new int} \langle S_n \rangle$ 
6: for each server  $\in S_n$  do
7:   if  $\text{server}[i] \text{ instanceof } (\text{vGPU})$  then
8:      $HQ = \max(\text{server}[i].\text{freeCore}/\text{configOf}(\text{vGPU}), \text{server}[i].\text{freeMem}/\text{configMem})$ 
9:   else  $SQ = \max(\text{server}[i].\text{freeCore}/\text{configOf}(\text{vCPU}), \text{server}[i].\text{freeMem}/\text{configMem})$ 
10:  end if
11:  if  $SQ[i] > 0$  then
12:     $NH_n.\text{push}[\text{server}]$ 
13:  end if
14:  if  $HQ[i] > 0$  then
15:     $H_n.\text{push}[\text{server}]$ 
16:  end if
17:   $R_s.\text{add}(\text{server}[i].\text{rack\_id})$ 
18: end for
19:  $R_{\text{quota}} = \text{new int} \langle R_s.\text{length}() \rangle$ 
20: for each rack  $\in R_{\text{quota}}$  do
21:    $\text{rack}[j] = \sum_i SQ[i] \forall i \text{ such that } \text{server}[i].\text{rack\_id} = j$ 
22: end for
23: sort  $NH_n$  and  $H_n$  by DESC
24: stable sort  $NH_n$  and  $H_n$  by DESC where  $R_{\text{quota}}[\text{server}[i].\text{rack\_id}]$ 
25:  $PP = \text{new vector} [\text{int}]$ 
26:  $PP_m = \text{new Map} \langle \text{String}, PP \rangle$ 
27: for each nonHPCHost  $\in NH_n$  do
28:   for each serverQuota  $\in SQ[PP[i]]$  do
29:      $PP_m.\text{push}(\text{vCPU}, PP[i])$ 
30:   end for
31: end for
32: return  $PP_m$ ;

```

- $S_n$ : number of servers
- $NH_n$ : number of Non HPC refined hosts
- $H_n$ : number of GPC based HPC hosts
- $R_s$ : set of racks
- $SQ$ : server quota
- $HQ$ : server HPC quota
- $PP_m$ : predictive policy map

Using this, we will have a scheduling policy having the rules where every rule will have the ordered list of all hosts of a rack on the basis of its quota of containers such that  $S_1, S_2, \dots, S_n$ . Here,  $S_1$  will have higher quota of containers in comparison to  $S_2$  and so on. The scheduler will place the containers in a rack with highest rackQuota, and the server in that rack will have highest serverQuota. On the top of this, based

**Table 1** Test bed specifications

Resource	Private cloud instance (VM)	Containers
Operating system	CentOS 7.0	CentOS 7.0
Memory	256 GB	126 GB
Network	Emulated 1GigE	Emulated 1GigE
Number of processors	12 X Intel Xeon X5650 @2.67 GHz	12 X QEMU virtual CPU @2.67 GHz

on user requirement, the matching configuration-based container will get picked and allocated to run application. To make sure of homogeneity, scheduler will refine the servers based on their processor type like GPU ones in one list and non-GPU will get clustered in different one. After refinement, scheduler policy will get applied and predicts the placement of container with best configuration. Currently, we have refined basis on GPU and CPU frequency.

## 4 Evaluation Methodology

This section covers and describes the experimental results and specification using our in-laboratory built cloud stack and the used application.

### 4.1 Experimental Test Bed

We have created our own cloud environment using a cluster of GPU and non-GPU hosts. We have provisioned the containers using Docker and manages our cluster with Kubernetes. Containers are the best choice for OS virtualization in cloud and to run HPC application. The containers are used for the analysis and experiments with different flavors of type minor 1 (2 core, 4 GB in terms of memory, 40 GB in terms of disk). We have used three servers which are having following configurations (Table 1):

### 4.2 Benchmarks and Applications

For benchmarking, we have chosen specific benchmarks and practical consideration of HPC application execution. Jacobi2D is a kernel generally runs 5-point stencil calculation to an average value in a 2D grid. This kind of stencil calculation is mostly used for research simulations in scientific area, image processing, and numeric

algebra. NAMD [7] is a scalable dynamics molecular application which represents a complex application that is being used on supercomputers. For our experiments, we have used the ApoA1 input (90 k atoms).

### 5 Results

At first, we have executed our proposed scheduling approach with predictive policy and evaluated the same. Using this, we have gained a performance on container placement matching the criteria of HPC-based application as per user-specific requirement. Ping pong benchmarking has been done by using a converse to compare bandwidth and latencies of container placement and found the virtualization overhead for the communication between containers. Using our predictive policy, compare any random scheduling and placement of containers of a cluster within same rack results into 8% of improvement. In cloud where core count is going to be high, it will really benefit in terms of scheduling computing instances. Figure 4 shows the effect of container placement based on underlying available servers using our predictive policy map having categorized host to choose either for HPC or normal applications in cloud. We have distinguished two cases—where eight containers are mapped under homogeneous environment (same type of processors) and four in heterogeneous (two are moderate processors where rest to the quite faster). We have computed the improvement as  $(CHetero - CHomo)/CHetero$  (Fig. 3).

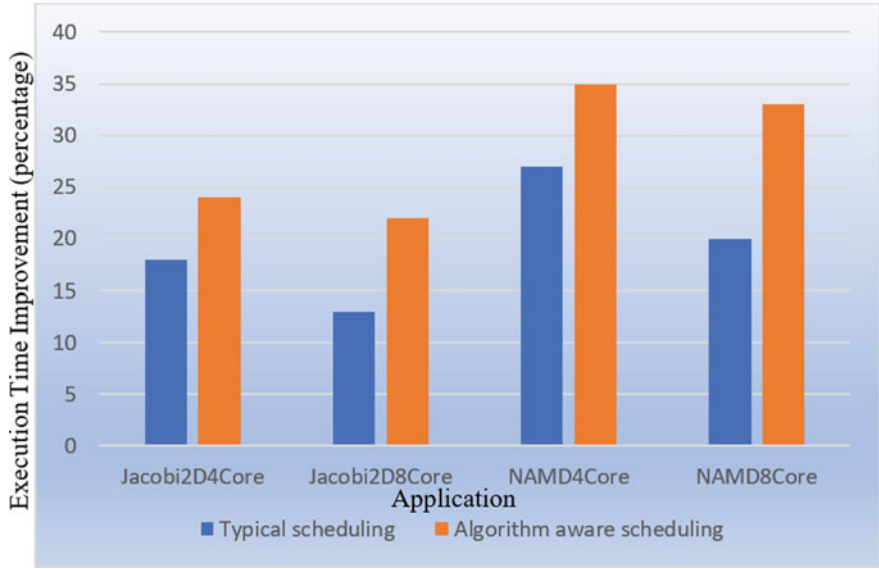
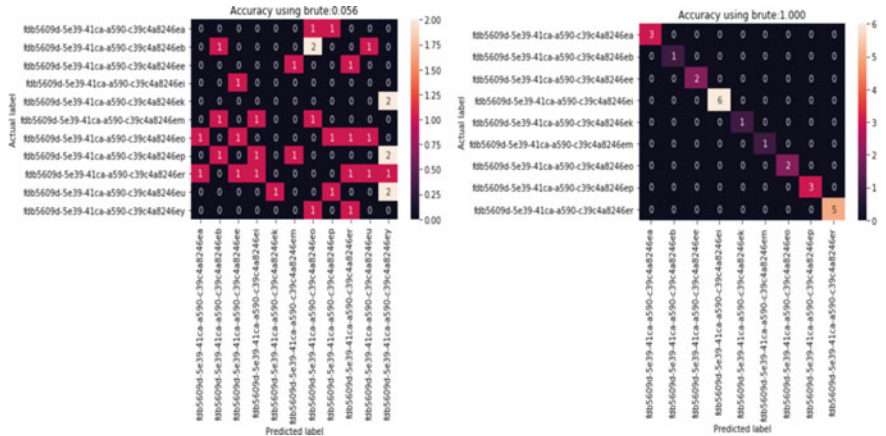


Fig. 3 Comparison between typical and algorithm-based scheduling



**Fig. 4** Containers placement analysis based on predictive policy map

It has been analyzed and found that the significant improvement is achieved based on application type and its workload. Since in HPC, parallel execution of application process time also includes the communication time in spite of processor speed, the improvement is indirectly proportional to the ratio of sequential execution time on moderate processor to that on quite faster one. We have gained approx. 22% improved execution time—which implies that we have save ~ 22% of execution time multiply by N CPU or GPU hours, where N is the total count of used processors.

We use the Docker for spawning the containers which really benefit us in world of virtualization. We were able to provision containers very fast even in distributed environment to achieve isolation. We have used the bridge networking here, i.e., default one. Kubernetes is mainly used to manage the clusters having containers. If one container crashes due to memory spike or CPU usage, it will automatically shift the application to a new one by launching the same within a few seconds.

## 6 Discussion and Conclusion

As HPC gains a popularity in cloud with the usage of containers, there are many research work on running HPC application in cloud computing with several benchmarking-like real applications which concluded on I/O performance [8], and interconnect channel interaction on enterprise cloud can downgrade the performance. Cloud is not completely a replacement to supercomputers in terms of metric \$/GFLOPS for horizontal scalability of HPC applications. It may be affecting the cost for few applications' execution on cloud rather than supercomputers, which needs less communication and low scalability. We demonstrated that if we to consider the container placement to run HPC application, platform will depend upon application requirement, its performance, and user configuration. We used Kubernetes for load



balancing system to manage the deployment of containers, self-healing, and monitoring. In our earlier work, we have demonstrated that how containers are more useful to run HPC application [9] instead of VM using Docker and Kubernetes. In this paper, we take next step and addressed the issue of container placement strategies which results significant improvement in performance. Our aim is to place the containers suitable to run HPC applications—which consist of  $k$  parallel instances typically requiring synchronization through, interprocess communication. Existing scheduling policy does not address this problem efficiently. Cloud computing management platforms or systems such as Open Stack [10] and Open Nebula [11] lack strategies that are totally based on the virtual machines usage instead of containers which are tightly coupled in nature for comprising one request for having scheduling decisions. We have proposed the GPU-based selection of container placement in case of HPC applications where are vCPU based to the normal ones. We have also considered the heterogeneity. Recently, there are many efforts are being put for job scheduling such as Load Sharing Facility (LSF) [12]—an enterprise job scheduler for load sharing via distributing jobs to available CPUs in heterogeneous network. SLURM, ALPS, MOAB, Torque, Open PBS, PBS Pro, SGE, and Condor are few other examples. However, they perform scheduling at the granularity of physical machines.

In this paper, we have concluded on the how part of predictive policy map usage, containers can be efficiently placed to execute HPC or non-HPC applications. We have created a scalable framework using Docker and Kubernetes by following our proposed scheduling approach and demonstrated the benefits of running HPC and simple application in cloud. Specifically, we focused on implementing the predictive policy to schedule containers in a cluster belongs to same rack instead of distributed one [13]. In future, we are planning to consider other factors also which can affect the container performance in terms of network and disk and can address other existing challenges to adopt HPC in cloud environment [14].

## References

1. Subhash S, Steve H, Haoqiang J, Johnny C, Robert H, Piyush M, Rupak B (2012) An application-based performance evaluation of NASA's Nebula cloud computing platform 336–343. <https://doi.org/10.1109/HPCC.2012.52>
2. Magellan: a cloud computing testbed. <https://www.nersc.gov/research-and-development/archives/magellan/>
3. Mohamed K, Muhammad H, Sameh S, Ahmed S, Ahmed EM (2015) Performance analysis of Eucalyptus Vs. EC2 Cloud for HPC scientific computing. <https://doi.org/10.13140/2.1.3883.0889>
4. What is Kubernetes, why you need Kubernetes and what it can do. <https://kubernetes.io/docs/concepts/overview/what-is-kubernetes/>
5. Gravvanis GA, Morrison JP, Marinescu DC et al (2018) Special section: towards high performance computing in the cloud. *J Supercomput* 74:527–529. <https://doi.org/10.1007/s11227-018-2241-9>
6. Amazon elastic compute cloud (Amazon EC2). <http://aws.amazon.com/ec2>
7. Scalable molecular dynamics with NAMD on the summit system. <https://www.ks.uiuc.edu/Research/namd/2.13/NAMD-IBM-Journal-Manuscript-Revised.pdf>

8. Arima E, Schulz M (2020) Pattern-Aware Staging for Hybrid Memory Systems. In: Sadayappan P, Chamberlain B, Juckeland G, Ltaief H (eds) High performance computing. ISC high performance 2020. Lecture notes in computer science, vol 12151. Springer, Cham
9. Abhishek MK, Rajeswara Rao D (2020) Dynamic allocation of high-performance computing resources. *Int J Adv Trends Comput Sci Eng* 9:3528–3543. <https://doi.org/10.30534/ijatcse/2020/159932020>
10. Pavle I, Harald R (2018) OpenStack cloud tuning for high performance computing. 142–146. <https://doi.org/10.1109/ICCCBDA.2018.8386502>
11. Eddy C, Lamiel T, Jonathan RC (2013) Comparison on OpenStack and OpenNebula performance to improve multi-Cloud architecture on cosmological simulation use case
12. Load sharing facility (LSF). <http://www.serc.iisc.ac.in/software/load-sharing-facility-lsf/>
13. Abhishek G, Dejan M, Laxmikant K (2012) Optimizing VM placement for HPC in the cloud, 1–6. <https://doi.org/10.1145/2378975.2378977>
14. Sachchidanand S, Nirmala (2016) Containers & Docker: emerging roles & future of cloud technology, 804–807. <https://doi.org/10.1109/ICATCCT.2016.7912109>

# Mango (*Mangifera Indica* L.) Classification Using Shape, Texture and Convolutional Neural Network Features



Sapan Naik, Yash Rana, and Viraj Thakkar

**Abstract** There are various qualities of foods which are grown in India. There is an important need to classify these fruits into grades/categories for the farmers to get optimum profits. However, this task is much manual and tedious. Thus, automation can help to get the task done. Here, mango (*Mangifera indica* L.) classification is performed. In this paper, we picked up seven categories of mangoes, i.e., Aafush, Kesar, Jamadar, Rajapuri, Totapuri, langdo, and Dahseri. First, we have prepared the dataset, and next shape, texture, and pretrained convolutional neural network (CNN) model's features are extracted and finally classification is performed using linear classifiers. Combination of size parameters and chain code are used for shape, gray level co-occurrence matrix (GLCM), scale-invariant features transform (SIFT) and local binary pattern (LBP) are used for texture identification. Five CNN models namely Inception v3, Xception, ResNet, DenseNet, and MobileNet are used for feature extraction while three linear classifiers support vector machine (SVM), multilayer perceptron neural network (MLP), and K-nearest neighbor (KNN) are used. In experiments, highest Rank-1 accuracy of 98.5% is achieved with features of Inception v3 Model and MLP classifier.

**Keywords** Convolutional neural network · Mango classification · Pretrained CNN models · GLCM · SIFT · LBP · Inception v3

## 1 Introduction

In the Indian economy, agriculture is the essential key sector which has played a vital role. Three major phases divide this production process into cultivation, harvesting, and post-harvesting. There is a wide scope of automation introduced in the field of agriculture like sensor, robots, computer, and machine vision technology. Scope of automation benefits the post-harvesting phase which have processes like cooling, cleaning, sorting, grading, and packing.

---

S. Naik (✉) · Y. Rana · V. Thakkar

Babu Madhav Institute of Information Technology, Uka Tarsadia University, Surat, Gujrat 394350, India

Our scope for post-harvesting phase is limited to automatic non-destructive fruit classification. It is identified through its parameters such as aroma, color, strength, composition, size, shape, texture, defects, and maturity [1]. Fruit classification is the pre-step of sorting and grading. In the field of agriculture, grading and classification is necessity. Manual classification and grading method are time-consuming, laborious, less efficient, monotonous, and inconsistent. In contrast to that, rapid, hygienic, consistent, and objective assessment are provided by automatic systems [2]. Reason for choosing mango fruit for the study is that it benefits us extraordinarily through its high-quality standards and ample of nutrients filled in it. Gujarat is a leading state with the largest area under mango cultivation stretching from Aafush, jamadar, Totapuri, Dasher, Neelam, Langdo, Kesar, Payri, and Rajapuri to Alphonso [3].

Nowadays, deep learning and Convolutional neural network (CNN) are used for image classification. After Krizhevsky and team [4] won ILSVRC2012—ImageNet classification competition, deep learning is being constantly improved. Studies in the initial models of CNN increased which resulted in the initiation of modern CNN architecture models such as Xception, MobileNet, Inception, and ResNet [5].

CNN can be implemented in four different ways for classification. First, pretrained CNN model can be used for classification where only final layer of CNN gets tuned and trained for new image dataset. Transfer learning technique is used here. Second, CNN model is used as feature extractor and extracted features are given to linear classifier for classification. Third, pretrained CNN model is fully trained from scratch and tested. Finally, one can create own CNN architecture model. Below is the review of the work that has been done for fruits and vegetable classification.

## 2 Related Work

Mangoes are classified using three parameters which are shape (measured in form of eccentricity, extent and cross-ratio), size (combination of area and weight), and maturity (the mean of channels— $a^*$  and  $b^*$  channels of  $L^*a^*b^*$  color model) in Naik et al. [2] where 90% accuracy is achieved. Mango is classified into six maturity stages in Mim et al. [6] using 24 image features (based on color features) and decision tree-based technique, where they have achieved 96% accuracy. 981 images of Alphonso mango are classified in 4 classes using  $L^*a^*b^*$  color space features and SVM as classifier in Raghavendra et al. [7] where 88% accuracy is achieved. Two types of mango fruits' images—thermal and RGB—are taken and classified mangoes into three categories in Bhole and Kumar [8]. Transfer learning is used with pretrained model, SqueezeNet, which gives 93.33% accuracy. Feedforward network trained with backpropagation neural network and the radial basis function network are compared in Olaniyi et al. [9] for mango grading in two classes with GLCM feature as input.

Late and Early fusion methods are explored for a mixture of NIR and RGB images as faster region-based CNN is used for fruit classification in Sa et al. [10]. Classification experiments are performed on seven fruits. Selective search algorithm and means of an entropy method are used for region extraction of fruit in Hou et al.

[11] and given as input to CNN where final classification is performed using voting mechanism. For online food materials classification, a fast auto-clean convolutional neural network-based approach is proposed in Chen et al. [12], which used adapting learning for multiclass prediction task. 13-layer convolutional neural network is used with three types of data augmentation method and comparison of max and average pooling in Zhang et al. [13] for fruit classification where authors have achieved 94.94% accuracy. Faster R-CNN is utilized for real-time multi-class fruit classification in Wan and Goudos [14] where framework contains creation of dataset, data augmentation, improved version of faster RCNN model generation and evaluating model performance. Pure convolutional neural network with seven convolutional layers and minimum parameters are introduced in Kausar et al. [15] and tested on fruit-360 dataset with accuracy of 98.88%. Handcrafted features in RGB color space, namely the centroid obtained from K-means clustering, the histogram, and a single color are added in CNN for making the stores' checkout process rapid in Rojas-Aranda et al. [16] where 95% accuracy is achieved for fruits with no bag and 93% for fruits in bag.

## **2.1 Contributions**

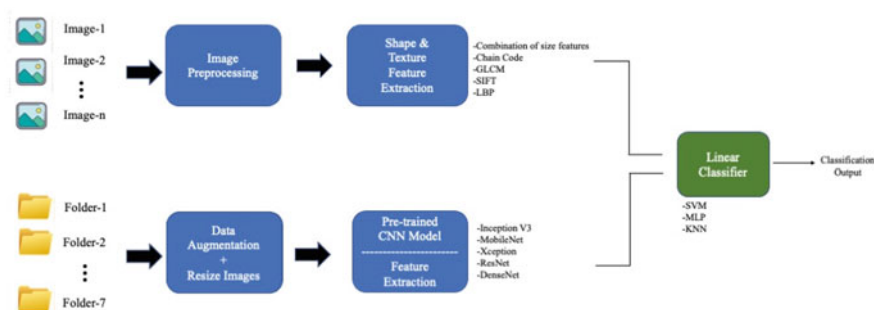
We have created a dataset of 7000 mango images of seven mango categories (Aafush, Kesar, Jamadar, Rajapuri, Totapuri, langdo, and Dahseri) where each category having 1000 images. Inception v3, Xception, DenseNet, resNet50, and MobileNet pretrained CNN models are used for feature extraction. For training and testing, traditional image features (shape and texture) are combined with CNN features. They are used with support vector machine (SVM), K-nearest neighbor (KNN), and multilayer perceptron neural network (MLP) classifiers.

## **3 Materials and Methods**

Proposed approach is divided into three phases, i.e., dataset preparation, feature extraction, and classification as shown in Fig. 1. Below we have discussed each phase.

### **3.1 Dataset Preparation**

The dataset of 7000 images were captured in daylight with a handheld camera covering the top-view of a mango (by placing the mango underneath the camera) and white paper as a backdrop. Figure 2 shows some sample images.



**Fig. 1** Procedure for mango classification using shape, texture, and CNN features



**Fig. 2** Sample images of mango dataset with Aafush, Kesar, Jamadar, Rajapuri, Totapuri, langdo, and Dahseri categories

### 3.2 Feature Extraction

Two tasks, i.e., shape and texture feature extraction and CNN feature extraction are performed parallelly as shown in Fig. 1.

*Shape and Texture Feature Extraction.* Physical non-destructive properties need to be studied to classify different categories of mango. By the study of seven categories of mango, it is concluded that all these mangoes are varied in shape and texture parameter. For extracting shape and texture parameters of mango, combination of size features, chain code, scale-invariant features transform (SIFT), local binary pattern (LBP), and gray level co-occurrence matrix (GLCM) features are applied. Here, first three features are considered for shape parameter while last two are for texture parameter [9, 17] because of their better performance compared to other available features. Firstly, input images are preprocessed. As a part of preprocessing, we have done basic noise removal and performed segmentation. Features are extracted after segmented images are received. The brief of features used is given below.

Fruit shape can be simply described using mathematical expressions related by global dimensions which is called combination of size features [17]. We have considered below size features.

*Circularity*: Relation  $(\text{Perimeter})^2/\text{Area}$  is used to determine how tightly packed an object's form is.

*Eccentricity*: measure of how much the deviation of the curve has occurred from the circularity of the given shape. The terms aspect ratio and major to minor axis ratio are almost equivalent to eccentricity.

*Bending energy*: states the energy needed to convert the particular shape under consideration into circle.

*Extent*: proportion of pixels in the given region to pixels in the total bounding box.

*Orientation*: Refer to the angle between the two axis (x-axis and major axis) of the ellipse created from an object under consideration.

Chain code is shape feature which is based on the concept of following an object's boundary and storing the direction of movement. The chain code is made up of straight-line segments that must all fit into an eight-position grid.

The SIFT algorithm takes an image and converts it into a series of features vectors that are invariant to the image's size, rotation, and translation, defining the image's local regions [18]. For SIFT algorithm, identification of scale-space extrema, keypoint localization, assigning orientation, and keypoint descriptor are the four core intuitions.

Haralick proposed a new method GLCM for texture feature extraction. It is derived by counting the occurrences of the pixel pairs occur in image  $f$  in the position  $Q$  specified by operators [19]. Some descriptors used for characterizing GLCM of size  $K \times K$  are: maximum probability, correlation, uniformity/energy, homogeneity, and entropy.

LBP is a quick and efficient texture descriptor that labels an image's pixels by quantizing each pixel's neighborhood and displaying the result as a binary number [20].

*CNN Feature Extraction*. There is requirement of fixed size input in CNN, captured images are resized from 2112\*4608 to 224\*224 and 299\*299 pixels according to CNN Model under used. This is divided into seven respective folders under their labels. Data augmentation technique is used to solve the issue of overfitting (using ImageDataGenerator method of TensorFlow). Horizontal flip, rotation, height and width shift, shear and zoom range are utilized for generating new images. Data augmentation is performed on each category to double the image. So, this results into total 2000 images per category and total of 14,000 images. For each category, 80% are chosen for CNN training, 20% for validation and 50 images of each category are used for testing. Inception v3, Xception, MobileNet, ResNet, and DenseNet architecture models of CNN are selected. The final completely connected layers of a pretrained CNN model are used to extract features, which are then saved locally in the HDF5 format. For experiment with all pretrained models, parameters are configured as.

Epoch value—1000,

Learning rate—0.01,  
Training batch size—100,  
Validation percentage—10.

3.3 Classification

Now, features extracted from above steps are given to different classifiers namely SVM, MLP (Initial learning rate for training: 0.001, maximum iteration:200, tolerance for optimization: 1e-4, momentum for gradient decent: 0.9, maximum epochs to meet tolerance:10) KNN. Classifiers are trained, validated, and tested. Below section summarizes the results.

4 Results and Discussion

Implementation is done in Python using Keras and TensorFlow libraries. Experiments are executed on macOS Big Sur (version: 11.2.3) in dual-core Intel Core i5 MacBook Pro (2017 model) with configuration of 8 GB RAM and Intel Iris Plus Graphics 640 card.

Shape features (combination of size feature and chain code) and texture features (SIFT, GLCM, and LBP) are extracted for each image and classifiers are trained for each image individually. CNN features are extracted for each image, and it is stored as HDF5 file. So, classifier is trained single time only for CNN features. Feature vector size of shape features is 9, while for texture feature is 266 and for all pretrained CNN models, it is 1000.

Table 1 shows the parameter tuning for KNN, SVM, and MLP. Table 2 summarizes the training time (feature extraction time) for each model. Table 3 summarizes overall result of all classifiers with different CNN architecture models.

It is observed from the experiment that MobileNet was fastest to get trained with lowest weight size, while DenseNet took longest time. The highest accuracy of 98.5% is achieved using MLP classifier with Inceptionv3 CNN model. KNN classifier does not perform well. All the pretrained model works fine as more than 94% accuracy is achieved in all the CNN models with SVM and MLP as classifiers. We have observed that handcrafted features play very vital role in achieving this accuracy. Combining

**Table 1** Parameters tuned for KNN, SVM, and MLP classifiers

Classifier	Parameters
SVM	Linear kernel, one vs rest, Tolerance:1e-3
MLP	Neurons:100, activation/optimizer function: relu/adam
KNN	Neighbors-5, algorithm-ball tree, distance metric-Minkowski



**Table 2** Feature extraction time and size of weights for Pretrained CNN model

CNN models	Time for feature extraction		Size of weights (MB)
	Start time	End time	
Inception V3	10:30	12:49	91.6
Mobile Net	11:30	13:30	16.40
ResNet	19:24	21:41	98.10
DenseNet	21:57	00:18	78.7
Xception	00:36	03:00	87.7

**Table 3** Classification accuracy with shape, texture and CNN features with linear classifiers

Classifier	Pretrained CNN models for CNN features				
	inception V3	mobilenet	Xception	ResNet	DenseNet
SVM	0.97	0.94	0.965	0.953	0.951
MLP	0.985	0.97	0.976	0.976	0.98
KNN	0.89	0.91	0.89	0.9053	0.92

shape and texture features improves our results by almost 4 to 5%. We have received 91.48% accuracy by only using CNN as feature extractor with linear classifiers [21].

By observing confusion matrix, we have seen that major misclassification occurs in Kesar and Rajapuri mangoes while other minor misclassification occurred in Lando and Totapuri categories of mango. Due similarity in shape, SIFT feature does not impact much overall accuracy, but texture features play an important role in classification. We have observed that GLCM performs better for texture feature extraction.

## 5 Conclusion

We have achieved highest accuracy of 98.5% and average of 94% accuracy for seven categories of mango. By observing and experimenting with handcrafted features, we can identify features which impact the most for classification and ignore the other features to reduce classification time. The above proposed method can be generalized for fruits of south Gujarat, India and a greater number of categories can be identified. Results can be improved and accuracy can be increased by fine tuning of parameters and combining more features and machine learning methods with CNN. There is a scope of decreasing time required to classify a mango. This classification can help in grading and detection of skin disease. From the above experiment, we can infer that CNN can very well be used for feature extraction, our linear classifier can be used for classification purposes. Moreover, chain code can be used for shape feature, and GLCM can perform very well as a texture feature.

## References

1. Slaughter DC (2009) Nondestructive maturity assessment methods for mango. University of California, Davis, pp 1–18
2. Naik S, Patel B, Pandey R (2015) Shape, size and maturity features extraction with fuzzy classifier for non-destructive mango (*Mangifera Indica* L., cv. Kesar) grading. In: 2015 IEEE technological innovation in ICT for agriculture and rural development (TIAR). IEEE, pp 1–7
3. Shah N, Patel C, Patel V, Attar S, Patel A (2013) Morphological description of mango varieties under agroclimatic conditions of Gujarat. AICRP (STF), CISH, Lucknow, India
4. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 25:1097–1105
5. Using Keras pre-trained deep learning models for your own dataset—Gogul Ilango. Available: <https://gogul09.github.io/software/flower-recognition-deep-learning>. Accessed 28 Jan 2019
6. Mim FS, Galib SM, Hasan MF, Jerin SA (2018) Automatic detection of mango ripening stages—An application of information technology to botany. *Sci Hortic* 237:156–163
7. Raghavendra A, Guru DS, Rao MK, Sumithra R (2020) Hierarchical approach for ripeness grading of mangoes. *Artif Intell Agric* 4:243–252
8. Bhole V, Kumar A (2020) Mango quality grading using deep learning technique: perspectives from agriculture and food industry. In: Proceedings of the 21st annual conference on information technology education, pp 180–186
9. Olaniyi EO, Oyedotun OK, Ogunlade CA, Khashman A (2019) In-line grading system for mango fruits using GLCM feature extraction and soft-computing techniques. *Int J Appl Pattern Recognit* 6(1):58–75
10. Sa I, Ge Z, Dayoub F, Upcroft B, Perez T, McCool C (2016) Deepfruits: a fruit detection system using deep neural networks. *Sensors* 16(8):1222
11. Hou L, Wu Q, Sun Q, Yang H, Li P (2016) Fruit recognition based on convolution neural network. In: 2016 12th international conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD). IEEE, pp 18–22
12. Chen H, Xu J, Xiao G, Wu Q, Zhang S (2018) Fast auto-clean CNN model for online prediction of food materials. *J Parallel Distrib Comput* 117:218–227
13. Zhang YD, Dong Z, Chen X, Jia W, Du S, Muhammad K, Wang SH (2019) Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools Appl* 78(3):3613–3632
14. Wan S, Goudos S (2020) Faster R-CNN for multi-class fruit detection using a robotic vision system. *Comput Netw* 168:107036
15. Kausar A, Sharif M, Park J, Shin, DR (2018) Pure-cnn: a framework for fruit images classification. In: 2018 international conference on computational science and computational intelligence (CSCI). IEEE, pp 404–408
16. Rojas-Aranda JL, Nunez-Varela JI, Cuevas-Tello JC, Rangel-Ramirez G (2020) Fruit classification for retail stores using deep learning. In: Mexican conference on pattern recognition. Springer, Cham, pp. 3–13
17. Moreda GP, Muñoz MA, Ruiz-Altisent M, Perdignes A (2012) Shape determination of horticultural produce using two-dimensional computer vision—a review. *J Food Eng* 108(2):245–261
18. Pérez DS, Bromberg F, Diaz CA (2017) Image classification for detection of winter grapevine buds in natural conditions using scale-invariant features transform, bag of features and support vector machines. *Comput Electron Agric* 135:81–95
19. Olaniyi EO, Adekunle AA, Odekuoye T, Khashman A (2017) Automatic system for grading banana using GLCM texture feature extraction and neural network arbitrations. *J Food Process Eng* 40(6):e12575

20. Muhammad G (2014) Automatic date fruit classification by using local texture descriptors and shape-size features. In: 2014 European modelling symposium. IEEE, pp 174–179
21. Naik S, Desai P (2021) Mango ( *Mangifera Indica* L . ) classification using convo- lutional neural network and linear classifiers. In: Proceeding—2021 Third International Conference Sustainable Computer (SUSCOM 2021), pp.1–9

# Modeling Human Tendencies for Password Guessing



Rohit Mutalik, Dhairya Chheda, Zeeshan Shaikh,  
and Dhanashree Toradmalle

**Abstract** Passwords form a predominant part of all authentication mechanisms and likely will not be replaced soon, primarily because of their ease of use and straightforwardness. Human-generated passwords are particularly susceptible to guessing attacks as a consequence of the limitation of precise recall and hence are not random. In this paper, we make use of the excellent expressive power of sequence modeling neural networks such as LSTMs and GRUs to effectively guess passwords as compared to other cutting-edge password guessing techniques like Markov models, JtR, and PCFGs. LSTM and GRU models were able to match about 55% of the testing set when evaluated on the Rockyou dataset. This proves that sequence modeling neural networks can effectively learn the distribution of real passwords from previously leaked datasets.

**Keywords** Long short-term memory · Gated recurrent unit · Generative adversarial networks · Password guessing · John the ripper · Hashcat · Password analysis · IWGAN

## 1 Introduction

Password-based authentication, even with its glaring deficiencies, shows no signs of being replaced owing to its excellent balance of security and ease of use. Human limitations to recall and password-based authentication mechanisms requiring precise

---

R. Mutalik (✉) · D. Chheda · Z. Shaikh · D. Toradmalle  
Shah and Anchor Kutchhi Engineering College, Mumbai, India  
e-mail: [rohit.mutalik@sakec.ac.in](mailto:rohit.mutalik@sakec.ac.in)

D. Chheda  
e-mail: [dhairya.chheda@sakec.ac.in](mailto:dhairya.chheda@sakec.ac.in)

Z. Shaikh  
e-mail: [zeeshan.shaikh@sakec.ac.in](mailto:zeeshan.shaikh@sakec.ac.in)

D. Toradmalle  
e-mail: [ghanashree.toradmalle@sakec.ac.in](mailto:ghanashree.toradmalle@sakec.ac.in)

recall lead to the selection of short [1] human-memorable passwords and their repeated use for authentication on multiple different security systems. Several web applications and security systems proactively tackle security risks arising from weak passwords by requiring users to use a combination of several different types of characters such as special characters and numerals and using password guessing tools such as Hashcat and JtR to find weak passwords.

Secure systems requiring passwords for authentication are susceptible to guessing attacks if security measures such as rate-limiting and salting are improperly implemented, as in the case of SolarWinds Breach [2]. Hackers can also hash probable passwords and search for matches and cracked hashes password hash database dumps found online. Learning the password set distribution allows us to sample the least probable password sequences and thus evade the aforementioned attacks instead of relying on static strength evaluation rules implemented by most websites.

In the case of offline guessing attacks, where the password  $p$  has to be recovered when the hash  $h$  is known, rainbow tables [3] can be used. Oechslin's rainbow table attack makes use of a reduction function  $R$  that performs a mapping between the hash function  $H$  and the set of passwords  $P$ . Alternate calls of  $H$  and  $R$  are used to generate a single chain. This method requires low disk storage space; however, it suffers from false alarm where  $h$  is not in the generated chain even when extended  $h$  matches with the endpoint of the chain. The password cracking performance of rainbow tables also heavily depends on the reduction function  $R$  chosen.

Traditional password guessing tools such as Hashcat and JtR utilize two approaches. A brute-force attack is the first solution, whereas a dictionary attack and variations are the second.

Brute-force attacks focus on constructing a candidate a password generated by combining several different types of characters while taking into consideration, the minimum and maximum lengths of the password and the set of all legal characters. The performance of such attacks has been greatly improved by optimizing the tools to run on graphical processing units (GPU); however, large space of legal characters and long passwords require an enormous amount of time and render such an approach unusable.

Dictionary-based attacks make use of attack dictionaries for password guessing. Previously disclosed password sets and a combination of words from the English dictionary and numbers to generate password guesses. Tools such as Hashcat [4] and JtR [5] utilize several mangling and concatenation rules to generate more candidate password guesses and expand the existing dictionary. Hashcat provides several built-in rulesets such as best64 which reasonably improves the performance of password cracking. These rulesets amplify the existing dictionary by changing character capitalization, appending or prepending numbers and adding special characters to an already existing password in the dictionary. However, such an attack is constrained by the distribution of leaked passwords and the size of the dictionary. This method also struggles with finding rulesets that improve the cracking performance since analyzing statistical patterns of tens of millions of passwords is tedious.

In this research, we suggest the use of excellent expressive power of neural networks to analyze existing statistical patterns in the leaked dataset and sample novel sequences from the modeled distribution to generate password guesses. Approaches that perform probabilistic analysis on the password set are also explored.

## 2 Background and Related Work

In this section, we describe other password guessing approaches that have been studied previously. The vulnerability of authentication systems to password guessing attacks is extremely conditional upon the type of attack vector as well as the configuration of the system itself. Guessing attacks are rendered useless for systems that lock the user out for a certain duration after a specific number of incorrect guessing attempts.

### 2.1 Probabilistic Approaches

Markov models were suggested by Narayanan et al. for password guessing suggesting that phonetic similarity of words play an important role in the memorability of passwords [6]. The Markov model estimates the likelihood of the next character occurring by taking previous background characters into account. Given a prefix of length  $(n - 1)$ , an  $n$ -gram Markov model predicts the likelihood of the  $n$ th character. As a result, for any string  $c_1, \dots, c_t$ , probability estimation is done as follows:

$$P(c_1, \dots, c_t) \approx P(c_1, \dots, c_{n-1}) \cdot \prod_{i=n}^t P(c_i | c_{i-n+1}, \dots, c_{i-1}) \quad (1)$$

Narayanan et al. also proposed a hybrid algorithm that generated passwords that matched the Markovian filter and are accepted by a finite automaton. However, finding suitable regular expressions by analyzing the password set containing millions of passwords is difficult. This work was subsequently improved by Deurmuth et al. [7] by generating passwords in descending order of probability thus minimizing password cracking time by trying the most probable guesses first.

Weir et al. [8] proposed probabilistic context-free grammars (PCFG) for generating password guesses. PCFGs treat passwords as having a certain grammatical structure and a set of terminals that fit into those structures. Such methods generated preterminal structures by substituting digit string and special string terminals into base structures in order of decreasing probability. Furthermore, most probable candidate password guesses are enumerated first by substituting alpha string terminals into preterminal structures sorted in descending order of probability. The probability of the generated password is calculated by multiplying individual probability of the

base structure, preterminal structure, and the substituted string. In the experiments, PCFG-based password guessing mechanism outperformed JtR operating in wordlist mode.

## 2.2 Deep Learning Approaches

The use of artificial neural networks to model the password set distribution was proposed by Melicher et al. to gauge the strength of human-generated passwords. FLA primarily aimed at evaluating a password's strength and the number of guesses needed to crack it on the client-side browser while keeping the model as small as possible for achieving sub-second latency [9].

PassGAN [10] was developed by Hitaj et al. which makes use of improved training of Wasserstein Generative Adversarial Network (IWGAN) to generate password samples with probability distribution  $p_g$  that is close to the real password distribution  $p_{\text{real}}$ . Gulrajani et al. [11] proposed IWGAN and introduced gradient penalty as a measure of enforcing 1-Lipschitz Continuity for stable training of WGAN [12]. The goal of generator **G** is to transform the original input noise vector  $z$  into password guesses using a series of residual blocks consisting of two 1D convolutions such that the discriminator is unable to differentiate between genuine and counterfeit samples. The discriminator network **D** aims to differentiate between the real passwords sampled from the dataset as well as the counterfeit examples generated by the generator network **G** and output a score estimating the realness of the generated password. The objective function of GAN [13] is given as follows:

$$\min_{\theta_G} \max_{\theta_D} \sum_{i=1}^n \log f(x_i; \theta_D) + \sum_{j=1}^n \log (1 - f(g(z_j; \theta_G); \theta_D)) \quad (2)$$

Further attempt to model the password set distribution using IWGAN was done by Nam et al. in [14]. Here, both the generator and the discriminator were recurrent neural networks. Nam et al. also experimented with a dual discriminator approach yielding slightly better results.

## 3 System Design

We explored a wide set of parameters to improve the cracking performance and approached password guessing as a sequence modeling task using long short-term memory (LSTM) and gated recurrent unit (GRU) cells.

### 3.1 Model Design

To avoid the recomputation of recurrent cell states of context characters while outputting the probabilities of the next character for passwords with common prefixes, we use a shared layer architecture where the training and inference models both share the same underlying dense and LSTM layers (see Fig. 2). This allows us to train the model by requiring the entire password but to predict the probability of next character in password sequence, only the previously computed states and last generated character are required. This reduces the amount of computation required for sampling passwords with common prefixes. The training model performs a forward pass over the network to generate the probability distribution over the token space for the next character in the sequence and then performs a backward pass to update weights of the shared layers based on the gradients obtained from loss calculation. The inference model performs only forward passes to generate probabilities during the password sampling process.

### 3.2 Model Prediction

Neural networks, like Markov models, are taught to anticipate the next character in the password given the state vectors representing all preceding characters. The network, given a sequence of characters, also models the probability of an END token denoting the end of the password. The generation of passwords starts by feeding the model a START token after which the model outputs the probabilities of being the first character over the entire token space. For example, the probability of password 'pwd' is calculated by using the start token (t) as a context and then question the model for the probability of having 'p' as the first character, then the probability of 'w' as the second character after 'p', then the probability of 'd' as the third character after 'pw' and finally the probability of entire password 'pwd' by questioning the model for the probability of end token after 'pwd'. Figure 1 shows the password sampling process.

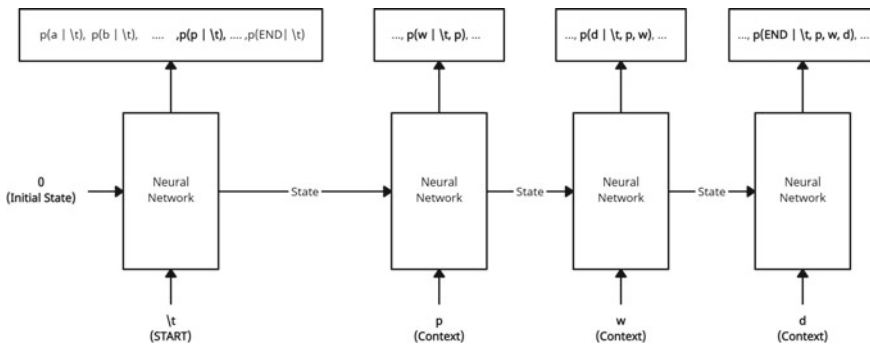
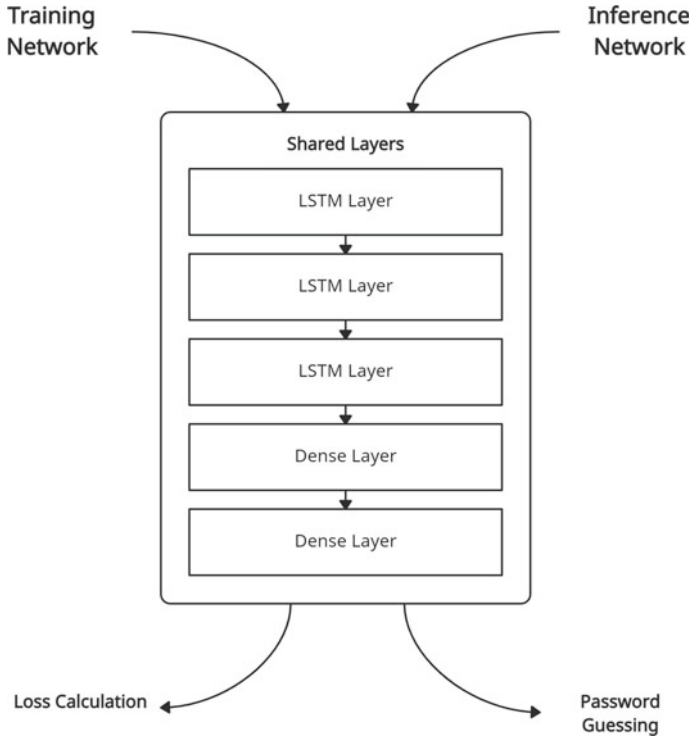


Fig. 1 Password sampling





**Fig. 2** Shared layer architecture

## 4 Our Approach

### 4.1 Token Space

In this work, we focus on a character-level password generation model including all characters, numerals, and special characters. To avoid straining the network by modeling all the characters, as shown empirically by Melicher [9], we choose to only model the probabilities of lowercase characters. Password guesses containing uppercase characters can be generated by post-processing the generated password to generate guesses consisting of both uppercase and lowercase characters based on the relative frequency of occurrence.

## 4.2 Model Architecture

We utilized the power of recurrent neural networks, which have been shown to process sequential data while maintaining internal state information about previously generated characters in the sequence. Recurrent neural network cells like long short-term memory [15] and gated recurrent units [16] which mitigate the issue of vanishing and exploding gradients and are excellent at generating text from the perspective of character-level language models.

## 4.3 Teacher Forcing and Exposure Bias

In this work, we implement teacher-forcing to train the neural network while taking into account both the internal memory state information and the previous timestep's ground truth context character. With this approach, the network can suffer from exposure bias during inference where the model can produce substandard sequences due to the propagation of errors. This happens because the network is trained to output conditional probability for the next character in the sequence, based only on the ground truth context character of the previous timestep and not on the generated errors.

## 4.4 Password Guess Sampling

To generate guesses using the inference model from the learned password set distribution, we use two approaches. First, we sample all the passwords that have the probability over a certain threshold. However, in this approach, the probability of the entire password reduces quickly as we are multiplying numbers that are less than one. This creates an undesirable effect where it prefers short passwords because of the fewer number of characters. In the second approach, we perform length normalization of the generated passwords and assign a score. This incentivizes the sampling algorithm to generate longer sequences especially considering the lack of passwords with a minimum of 16 characters in the Rockyou dataset. Passwords having a score greater than the threshold are then sampled. The normalization constant  $\alpha$  can be tuned as per the performance of the sampling algorithm.

$$\text{score} = \frac{1}{T_y^\alpha} \sum_{t=1}^{T_y} \log P(y^{<t>} | \text{START}, y^{<1>}, \dots, y^{<t-1>}) \quad (3)$$

$T_y$  denotes the length of the password.

## 5 Experimental Setup

### 5.1 Model Implementation

Our implementation makes use of the Keras Library with TensorFlow 2.3 [17] back-end for training both the networks. The training models used three stacked recurrent layers with 300 units followed by two densely connected layers with a total of 4,861,207 parameters. The training and sampling of passwords were done in Python 3.8. The tests were carried out on a computer running Ubuntu 20.04 LTS, a 6 core AMD Ryzen 5 3500 CPU, 16GB of RAM, and an NVIDIA GeForce 2060 Super extreme edition GPU with 8GB of VRAM.

### 5.2 Training and Testing Configuration

**Rockyou Dataset** To evaluate the performance of the LSTM and GRU networks and generate candidate password guesses, we trained these models on sets of passwords obtained from the Rockyou password dataset leak [18]. We believe this approach is innocuous, even though the passwords might contain personal information since the dataset is publicly available.

In this experiment, we selected all passwords of length less than or equal to 32 from the Rockyou dataset (about 14 million passwords). Before training the model, as a standard preprocessing step, all the unicode and special ASCII control and extended ASCII characters were removed. Out of the total 14 million passwords, about 70% (9,973,206 passwords) were used to train the LSTM and GRU networks. For evaluation purposes, we calculated the set difference between the training set and the remaining 30% of the passwords (4,126,994 total passwords and 3,921,227 unique passwords) resulting in 3,665,812 passwords that the models have not examined during the training.

**Candidate Password Sampling Configuration** Once the models are trained, the sampling algorithm generates password guesses with a score greater than the configured threshold. The score threshold and alpha parameters for normalization are mentioned in Table 1. The generated passwords are then post-processed using the steps mentioned in Token Space.

**Table 1** Sampling algorithm configuration

Model	Threshold	$\alpha$
LSTM	−8.465	0.37
GRU	−9.45	0.3275

6 Evaluation

6.1 Evaluating Generated Passwords

To evaluate the performance of the models, we generated about  $1.57 \times 10^9$  and  $1.53 \times 10^9$  guesses using the LSTM and GRU model, respectively. Inspection of the generated passwords shows that the model has learned to construct meaningful candidate passwords containing names and a combination of names and numbers such as name followed by 2 or 4 digits indicating the birth year of the user. Our results show that neural networks exhibit exceptional password guessing performance for long and complex passwords. For the evaluation of both the models, the entire testing set was further partitioned into three sets based on the length (1 to 8, 9 to 15, and 16 to 32 characters). Figure 3 highlights the relation between the number of password guesses and the total number of passwords matches across all sets for both the models. Our results show that the GRU model slightly outperformed the LSTM model in the  $\leq 8$  character testing set with 56% matches as compared to the LSTM model’s 50% matches out of the total 2,250,012 testing passwords. However, for the remaining two sets (9–15 and 16–32) LSTM performs better with 34% and 25% matches, respectively, as compared to the GRU model’s 32% and 15% out of the total 1,345,062 and 70,738 passwords in 9–15 and 16–32 sets. The findings of the evaluation are detailed in Table 2.

Both the models generate novel sequences by combining words, numbers, and specific frequently occurring substrings of passwords in the training set such as ‘love’ or ‘123’. For instance, the model generated ‘loveforever’, ‘love4ever’, and

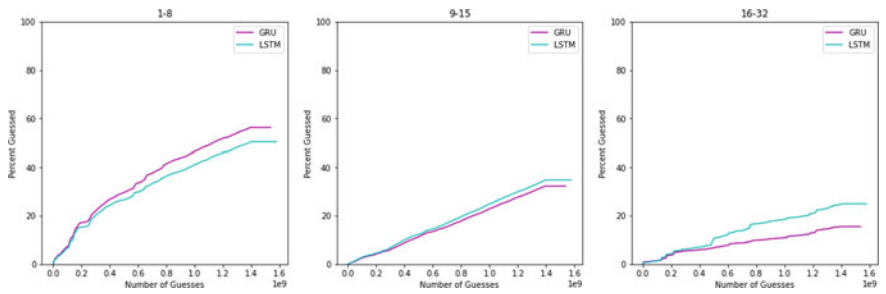


Fig. 3 1–8, 9–15 and 16–32 characters cracking performance

Table 2 Guessing performance of models

Password sets	≤8	≥9 and ≤15	≥16 and ≤32	Total
GRU model	1,266,478 (56%)	433,656 (32%)	10,986 (15%)	1,711,120 (46%)
LSTM model	1,133,612 (50%)	468,565 (34%)	17,628 (25%)	1,619,805 (44%)
Testing set	2,250,012	1,345,062	70,738	3,665,812

**Table 3** Sample non-matched passwords

ricardoisthebest	ihatemywife123	ashley1869
@ngelina	idontlikemybabyboy	27/09/06
ilovemyprincess70	strawberrytinkerbelle	nokia700
assholeisthebomb	christopher2007@yahoo.com	samurai@123
ilove4everross	teamomuchoasantonio	25aug2007

‘love4eva’, thus effectively capturing human password generation processes. This enables the models to investigate passwords in areas where conventional cutting-edge rule-based tools cannot. According to our research, the number of cumulative password matches gradually increases with increasing guess attempts. Both the models were able to match more than 40% of the  $\leq 8$  category passwords at just  $10^9$  guesses. For targeted password guessing approaches, tools such as common user password profiler (CUPP) can be used to make informed guesses, using commonly seen patterns, Personally Identifiable Information (PII) about the target, such as full name, date of birth, and nickname.

## 6.2 Non-matched Password Guesses

Closer examination of sampled passwords that were absent in testing set leads us to conclude that these passwords look appropriate for passwords generated by humans. We also believe that, while these passwords did not match with the testing set, they could still match with the passwords chosen by users on websites other than Rockyou. For targeted guessing, the generated passwords can be utilized as templates by replacing names and numbers in the guesses. Although carefully inspecting each generated password for producing targeted guesses is a tedious process, Table 3 lists a small sample of non-matched passwords.

## 7 Conclusion

We presented two neural networks with LSTM and GRU cells in this paper to learn the distribution of passwords chosen by humans using recently leaked password sets. These models were able to generate passwords similar to humans by exploring a significant portion of the password space without requiring any prerequisite knowledge and analysis of the leaked password dataset. These models have password cracking performance comparable to the Hashcat and JtR, which are cutting-edge password guessing tools even at lower training times.

The models were evaluated on the portions of the dataset that the networks were not trained on, thus determining how well the models were able to figure out how actual passwords are distributed. Our findings reveal that the models were able to generate human-like passwords and matched about 50% of the test set in just  $10^9$  guesses. The learned distribution can also be used to determine the strength of passwords selected by the user. Suggesting less probable and pronounceable passwords for users by utilizing the learned patterns from leaked passwords remains a high potential research area.

The ubiquitous nature of password-based authentication systems requires that developers adopt secure programming practices and implement password security measures such as two-factor authentication, salting, and limiting the number of incorrect password attempts.

## References

1. Morris R, Thomson K, Laboratories B (1979) Password security a case study
2. <https://www.zdnet.com/article/cisa-solarwinds-hackers-also-used-password-guessing-to-breach-targets/SolarWindsBreach>
3. Oechslin P “Making a Faster Cryptanalytic Time-Memory Trade-Off”, Laboratoire de Sécurité et de Cryptographie (LASEC), Ecole Polytechnique Fédérale de Lausanne, Faculté I&C, 1015 Lausanne, Switzerland
4. <https://hashcat.net/hashcat/Hashcat>
5. <https://www.openwall.com/john/> John The Ripper
6. Narayan A, Shmatikov V (2005) Fast dictionary attacks on passwords using time-space tradeoff. The University of Texas at Austin
7. Duermuth M, Angelstorf F, Castelluccia C, Perito D, Chaabane A (2015) OMEN: faster password guessing using an ordered markov enumerator. In: International symposium on engineering secure software and systems, Mar 2015, Milan, Italy, hal-01112124
8. Weir M, Aggarwal S, de Medeiros B, Glodek B (2009) Password cracking using probabilistic context-free grammars. Department of Computer Science, Florida State University, Tallahassee, Florida 32306, USA (2009)
9. Melicher W, Segreti SM, Komanduri S, Bauer L, Christin N, Cranor LF, Fast, lean, and accurate: modeling password guessability using neural networks. Carnegie Mellon University
10. Hitaj B, Gasti P, Ateniese G, Perez-Cruz F PassGAN: a deep learning approach for password guessing. In: Stevens Institute of Technology, New York Institute of Technology, Stevens Institute of Technology, Swiss Data Science Center (ETH Zurich and EPFL)
11. Ishaan G, Faruk A, Martin A, Vincent D, Aaron C (2017) Improved training of wasserstein GANs. montreal institute for learning algorithms, Courant Institute of Mathematical Sciences
12. Arjovsky M, Chintala S, Bottou L Wasserstein GAN. Courant Institute of Mathematical Sciences, Facebook AI Research
13. Goodfellow IJ, Pouget-Abadie J, Mirza M, Ozair S, Courville A, Bengio Y Generative adversarial nets. Département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal, QC H3C 3J7
14. Sungyup N, Seungho J, Jongsub M (2020) A new password cracking model. Korea University, Seoul, South Korea, Graduate School of Information Security
15. Hochreiter S, Schmidhuber J Long Short-Term Memory. Technische Universität München 80290 München, Germany, Corso, Elvezia 36 6900 Lugano, Switzerland

16. Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Bengio Y Learning phrase representations using RNN encoder-decoder for statistical machine translation. Université de Montréal, Jacobs University, Germany, Université du Maine, France, Université de Montréal, CIFAR Senior Fellow
17. <https://github.com/tensorflow/tensorflow/> Tensorflow Github Repository
18. <https://wiki.skullsecurity.org/Passwords> Rockyou Dataset

# Leveraging Big Data Analytics for Predictive Policing in India



Kashika Wanchoo and Jaimin Desai

**Abstract** India ranks 67th out of 135 countries on the Crime Index. (Numbeo, Crime Index by Country, 2021, [https://www.numbeo.com/crime/rankings\\_by\\_country.jsp?title=202](https://www.numbeo.com/crime/rankings_by_country.jsp?title=202)) According to the Indian Penal Code, there are almost 32.2 lakh criminal offences in India and 19.4 lakh in Special and Local Laws (SLL). As of 2019, this makes the safety of civilians very daunting. (Tom Reding, Jan 8, 2021 Crime in India, [https://en.wikipedia.org/wiki/Crime\\_in\\_India](https://en.wikipedia.org/wiki/Crime_in_India)) Police officers cannot handle crime in such a way that allows them to examine each case individually. They must be steered to felony-intensive areas pre-emptively. Big data is an area that deals with methods for analysing, systematically extracting information from and dealing with data sets that are too large or complex for conventional data-processing application software to handle. Over the years, there have been several implementations that have demonstrated the potential ability to leverage big data technologies. The paper focuses on predictive policing as one such implementation. The goal of the paper is to examine how big data can be applied to this segment of Indian government to locate inefficiencies and identify ways of minimizing them.

**Keywords** Big data · Indian government · Predictive policing · Hot spot analysis · Data analytics · Crime score

## 1 Introduction

The predictive policing concept is based on using case data to calculate future crimes to optimize police force allocation. Thus, predictive policing is based on various statistics and/or social research aspects. Predictive policing is often associated with the word and subject area big data, which is about the possibility of analysing a mass amount of data to provide immediate actionable results. By using the data, appropriate police measures can be determined to prevent potential crimes (preventive measures) or to investigate them more fully (repressive measures). This will allow police forces to deploy more efficiently. The aim of “predictive policing” is therefore to identify

---

K. Wanchoo (✉) · J. Desai  
NMIMS University, Mumbai, India



emerging hot spots based on a variety of crime factors and thus to achieve strategic and tailored police work [3].

### ***1.1 Predictive Policing in India: Current Scenario***

Recent changes in Indian law enforcement agencies have led several of them to adopt predictive analytics, which is a newer perception. Several Indian military units in addition to the Indian Intelligence Community are using predictive analytics to enhance their performance. India is particularly important for studying such deployments at this point since the country is presently engaged in efforts to deploy various machine learning technologies in all aspects of public life.

Predictive policing plans appear to be in the works, but first steps need to be taken in India that include digitizing data collection and connecting law enforcement agencies [4].

As an example, the law enforcement officers in Jharkhand now have a Naxal Information System, a Crime Criminal Information System (which will be blended with CCTNS) and a Geospatial Information System (GIS) that provide tailored plots for Maoist operations.

Officers from Jharkhand can access the Crime Analytics Dashboard from anywhere for real-time crime statistics and other crime statistics, including the number of crimes and their location, providing them with the real-time, up-to-date data they need to refine their awareness [5].

### ***1.2 COVID Perspective***

In this era of unprecedented global pandemics that are affecting all sectors of human life, India scores extremely high in the number of crises. The police force in the country is adversely affected by the same. Therefore, in addition to the work and maintenance of the current pandemic problems, there has been a shortage of manpower within the country and a delayed set of actions for some crimes that occur within the country.

COVID-19 will have a great effect on many aspects of society even after a vaccine is developed for it. It may alter the way we go to school and work. It might alter how health care is provided. But it will also alter patterns of criminal behaviour.

Criminal activity involving different crime types will change dramatically due to changes in time, location and mix of crimes. Without advanced analytics, it will be almost impossible to glean these changes in real time.

Departments, however, will be able to recognize and respond to the new patterns using predictive policing's machine learning algorithm, as well as changing their patrol patterns [6].

## 2 Data Preparation

### 2.1 Data Collection

In the first phase, the data can be obtained from a variety of sources, including crime statistics derived from police station databases, environmental information, such as crime patterns throughout the year, local neighbourhood composition and call logs from a mobile phone. In the following step, it aims to combine data from a variety of sources, including historical electronic data, physical copies of case files, information from current registers, papers, images and case files with pictures, FIRs, indictment records, memorandum of seizures, memorandum of conviction, memorandum of arrest, data from crime cases and incidents, offenders' files and police department records data [7]. Additional sources of valuable information may include The Popular Integrated Police Application, which was intended to be set up within the police departments to facilitate reporting and investigating crime, and prosecution. A sample data set is presented below with variables as follows (Fig. 1; Table 1).

Report Number	Date	Occur Date	Occur Time	Location	Shift Occurrence	Ward	Crime Reports	State	Latitude	Longitude	Crime Category
90010930	02-02-2021	01-03-2021	1145	Andheri East	Day	K/E	LARCENY-NON VEHICLE	Maharashtra	19.121090	72.868162	4
90011083	16-02-2021	04-03-2021	1330	Vikhroli	Day	S	LARCENY-NON VEHICLE	Maharashtra	19.118648	72.937051	4
90011208	23-03-2021	04-03-2021	1600	Dadar	Evening	G/N	LARCENY-NON VEHICLE	Maharashtra	19.019848	72.844580	4
90011248	04-03-2021	04-03-2021	2130	Malad	Evening	P/N	HOMICIDE	Maharashtra	19.187448	72.841455	1
90011250	13-03-2021	13-03-2021	330	Matunga	Morning	F/N	HOMICIDE	Maharashtra	19.027687	72.855029	1

**Fig. 1** Example of our sample data set

**Table 1** Variable description of the sample data set

Variable	Description
Report number	Case number
Date	Date of crime reporting
Occur date	Date at which crime occurred
Occur time	Specifies the time at which the crime occurred
Location	Area/Suburb where crime took place
Shift occurrence	Time of the day during offence
Ward	Ward branch of crime reporting
Crime reports	Type of crime reported
State	State of the country
Latitude and longitude	Geographical coordinates
Crime category	Category of crime

## 2.2 Utilizing Big Data Concepts for Implementation

**Data pre-processing:** There will be a vast amount of data in the crime data set that will help us learn more about crimes that have occurred over the years and provide us with detail regarding the crimes committed in each region. This data set can then be converted into a csv file that can be loaded into a structured database using the MySQL Query Browser.

**Module for data migration with Sqoop:** Data is transferred between Hadoop (HDFS) and relational databases using Sqoop. Sqoop can be used to perform a wide range of tasks. With the help of Sqoop tools, tasks such as fetching a particular column or fetching a set of columns that meet a certain condition can be handled easily.

**Module for data analysis in Hive:** The purpose of this module is to introduce us to Hive Query Language (HQL), a language that allows us to analyse data sets. The Hive framework allows us to build tables, joins, buckets and partitions as well as partitioning concepts.

**Partitioning and bucketing:** Partitioning is a more effective way to query the data. Partitioning easily lets us query one part of the table by grouping different columns together. This creates a multidimensional structure on directory storage. Data can be divided into manageable chunks called buckets or clusters and can be based on an algorithm hash function, which will use different algorithms depending on what column is being bucketed. The aim of this strategy is to distribute data evenly to save space.

**Data analytic module with MapReduce:**

Map stage: During the mapping stage, the mapper will process the input file to create several small chunks of data based on the input line.

Reduce stage: Data is read from the mappers using the reducer stage. Once the process has been completed, the results are saved to HDFS.

*MapReduce Process:*

- (i) **Splitting:** The inputs of a MapReduce job are split into discrete units referred to as input splits.
- (ii) **Mapping:** After obtaining input data and passing it to the mapping function, output values are generated. Examples include counting the number of times each crime occurred in input splits and turning it into a list of (crime, total) from input splits.
- (iii) **Shuffle:** A mapping phase produces this result. Specifically, it aims to converge the relevant records of the mapping phase.
- (iv) **Reducing:** A single output value is calculated by merging the results of the shuffling phase into a single output value [8].

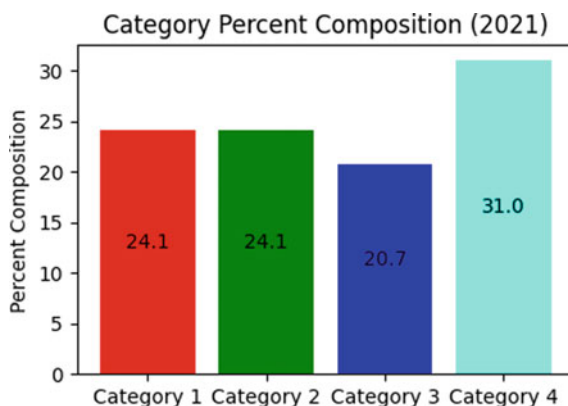
## 2.3 Data Analysis

The next step is analysis, which includes using a predictive algorithm to evaluate the data collected. Regression models such as the near-repeat theory posits that crime can lead to increased risk in nearby areas in the days following the original incident [8]. Another example is social network analysis (SNA) wherein an investigation of criminal networks, offending patterns and victimization can be derived. SNA offers a way to understand criminal behaviour through relationships and interactions [10]. These are some instances of proposals that can be utilized to gain information about an offender's activities. For example, PredPol makes use of these trends to notify its models about hot spots in a particular zone. For example, if the research reveals information about the causes and patterns of crime, such as seasonality and trends, the algorithms could identify hot spots in each region and concentrate their efforts on those patterns. In place-based theories, geographic location is emphasized as a contributing factor to criminal activity [7]. For example, PredPol forecasts where and when a crime is most likely to occur by gathering significant crime data from the police bureau and breaking it down into 500-foot-by-500-foot figures on a plot (Fig. 2).

## 3 Methodology

Predictive policing is based on the premise that crime is not spread uniformly through individuals or areas. As a result, crime trends are a “product of underlying environmental conditions that build vulnerability in victims and spaces at specific periods.” Thus, predictive policing attempts to make use of insights gained from algorithms rather than relying on hunches from police officers, by using the power of big data. More importantly, the application of algorithms also shows that there is a scope of reducing the biases which exist in decision-making. Predictive policing functions

**Fig. 2** Percentage composition of crime categories



as part of the global trend towards intelligent decision-making, and it interacts well with India’s recent initiatives to seek a data-rich jurisdiction (Table 2; Fig. 3)

As part of our methodology, each neighbourhood would be assigned a crime score for each day. Our strategy is to aggregate a score which considers the severity of different crimes with respect to these locations to label them as hot spots. These features will measure the level of crime in an area or for a specific day. Through the help of this crime score, as well as our predictor feature, we can come as close as possible to the use case of aiding law enforcement in determining crime across time and neighbourhoods. Data gathered can be classified into different categories based upon the severity of the crime. As a result, each category can be given a weight of  $1\times$ , with an increase in weight of  $10\times$  and so on for more serious crimes. Crime scores could be calculated by summing the categories of crime at different locations at certain time [11]. Following instance demonstrates how a crime score is calculated:

**Crime Score Calculation = (1000\* Category 1 counts) + (100\* Category 2 counts) + (10\* Category 3 counts) + (1\* Category 4 counts).**

**Table 2** Weighting crime by severity

Category	Crimes in category	Weight
1	Manslaughter, homicides	1000×
2	Robbery, aggravated assault	100×
3	Vehicle-theft burglary	10×
4	Larceny	1×

**Fig. 3** Locationwise mean crime score

Location	Average Crime Score
Vikroli	1
Dadar	3
Andheri East	5
Bandra West	10
Colaba	10
Dahisar	10
Goregaon	300
Kandivali	400
Matunga	1000
Andheri West	3000
Malad	3030



Fig. 4 Hot spot zones of crime

3.1 Geospatial Visualization of Hot spot Areas

Insights can be gathered from Fig. 4 that displays which areas are being affected by which types of criminal activity and how patrolling those areas can minimize them.

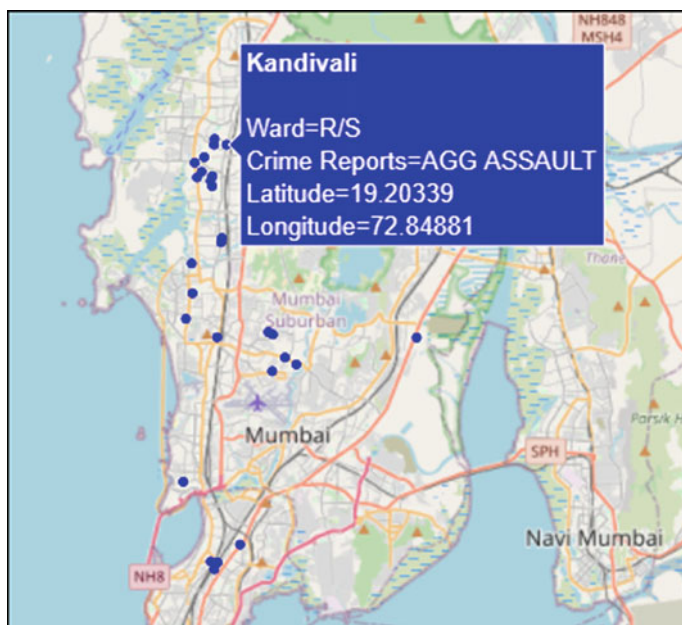
We can identify clusters at a location from individual data points as shown in Fig. 5.

4 Benefits

Statistical and analytical tools provide a variety of useful features for law enforcement forces deploying them.

4.1 Optimized Allocation of Resources

Police officers in India have been overburdened creating health and social issues. Due to the hectic working hours of police officials, technology or methods that increase the productivity of resource allocation are highly recommended. Since the law enforcement officials often work seven days a week, long hours are often mandatory.



**Fig. 5** Criminal activity by individual data points

## 4.2 Data Analysis

Utilizing predictive policing, authorities can take preventative measures against crime by focusing on crime-prone areas or individuals most likely to commit offences. Keeping criminals from becoming victims and offenders from making life-altering errors is a social and economic benefit for all associations, not just the victims who are at risk but also the offenders.

## 4.3 Better Informed Decision-Making

A predictive policing approach seeks to extract information from a diverse range of sources such as criminology, figures, actuarial science and the context of the surrounding area. Thus, a more holistic analysis of crime patterns is expected to be produced [7].

## 5 Drawbacks

Despite the significant success achieved by several agencies in different countries, some caution that through the uninhibited advancement towards policing through objective analysis, attention is being diverted from the challenges that can only be addressed through other means.

### 5.1 *Inadequacy in Curbing Abuses of Power*

Predictive policing is the pre-emptive response to the inherent crime risk. This automatically raises concerns with regard to the possibility of preventive measures hindering and interfering with the rights of innocent people. India's Code of Criminal Procedure contains provisions that permit arrests on suspicion. Abusing predictive policing may result in warrantless arrests and detentions without probable cause.

### 5.2 *Biased Learning Data*

The likelihood of existing inequities being amplified when data is used in decision-making increases. In addition, each predictive police action adds to the data that is used to make decisions about what to do next. Therefore, if the neighbourhoods are known to be crime-prone, the police will direct more attention to those neighbourhoods and consequently, they will receive greater policing resources. A particular form of discrimination in predictive algorithms is institutional due to implicit biases present in data. Several issues of concern arise regarding police data being used as 'history sheets' and 'rowdy sheets', as well as records of people who are classified as 'hooligans', 'goondas' and 'criminal tribes'. Furthermore, courts are unable to comprehend the predictive policing algorithms in many circumstances. As a result, despite striving to prevent injustice through rights guaranteed by the constitution striving to prevent injustice, redress mechanisms fail.

### 5.3 *A Data-Oriented Ideology*

Some jurisdictions have expressed concerns about relying too much on data and overlooking other aspects when policing. For instance, police officers could be more inclined towards only patrolling areas that are in the spotlight or highlighted on heat maps [7].



## 6 Conclusion

A comprehensive decision support system that utilizes big data analytics is critical for the successful resolution of increasingly complex crimes by government and judicial agencies on a regular basis. In the changing societal dynamics, where the number of people and the level of crime increase, it is crucial to acknowledge the large crime data sets to identify trends. In this paper, we have implemented a solution on a sample data set which can be used to develop strategies to predict hot spot areas with prime crime locations.

## 7 Future Scope

In the future, we plan to combine Indian demographic information with other databases. We can also improve our model by adding more features and more variability in locations so that it learns and becomes more accurate. While most police departments have implemented operational human resource management systems, efforts should be made to mine the data more effectively. A gap in the organization can be identified by analysing qualifications, age, gender, place of residence, caste, level of education, training, posting, rank and supervisor–employee ratios. We can address these gaps via hiring, training and postings to ensure a more ‘balanced and effective’ workplace. Likewise, officer performance can be scored based on several key performance indicators, including the numbers of crimes solved, the time taken to deal with complaints and the scores received from citizens.

**Acknowledgements** In appreciation, we wish to thank Professor Satyam Priyadarshy and Professor Sarada Samantaray for their patient assistance, enthusiastic support and critiques of this work, as well as for their assistance in keeping our work on track.

## References

1. Numbeo (2021) Crime index by country. [https://www.numbeo.com/crime/rankings\\_by\\_country.jsp?title=2021](https://www.numbeo.com/crime/rankings_by_country.jsp?title=2021)
2. Reding T (2021) Crime in India. [https://en.wikipedia.org/wiki/Crime\\_in\\_India](https://en.wikipedia.org/wiki/Crime_in_India)
3. Momsen C, Rennert C, Attorney at law (2020) Big data-based predictive policing and the changing nature of criminal justice. KriPoZ 3 | 2020—Criminal Political Journal
4. Benefits of predictive policing para 4 Available: <https://www.innefu.com/blog/benefits-of-predictive-policing/>
5. George R (2015) Predictive policing: What is it, how it works, and its legal implications. <https://cis-india.org/internet-governance/blog/predictive-policing-what-is-it-how-it-works-and-it-legal-implications>
6. Policing in the Time of Covid-19 <https://blog.predpol.com/policing-in-the-time-of-covid-19>. 26 Mar 2020

7. Sinha A (2017) Big data in Governance in India: case studies. <https://cis-india.org/internet-governance/files/big-data-compilation.pdf>
8. Nadathur AS, Narayanan G, Ravichandran I, Srividhya S, Kayalvizhi J (2018) Crime analysis and prediction using big data 119(12):207–211. <http://www.ijpam.eu>
9. Polans D (2018) Social network analysis for law enforcement. [https://crimegunintelcenters.org/wp-content/uploads/2018/07/iacawp\\_2018\\_02\\_social\\_network\\_analysis.pdf](https://crimegunintelcenters.org/wp-content/uploads/2018/07/iacawp_2018_02_social_network_analysis.pdf). Accessed 16 May 2021
10. Johnson SD, Bowers KJ (2014) Near repeats and crime forecasting. In: Bruinsma G, Weisburd D (eds) Encyclopedia of criminology and criminal Justice. Springer, New York, NY. [https://doi.org/10.1007/978-1-4614-5690-2\\_210](https://doi.org/10.1007/978-1-4614-5690-2_210)
11. Prasada S (2019) Predictive Atlanta crime mapping. <https://github.com/sanjeevprasada/AtlantaCrime>

**Kashika Wanchoo** is with NMIMS's Mukesh Patel School of Technology Management & Engineering, Mumbai, India.

**Jaimin Desai** is with NMIMS's Mukesh Patel School of Technology Management & Engineering, Mumbai, India.

# An Assured Encryption Strategy for Private Data Protection in Hybrid Cloud



D. I. George Amalarethinam and J. Madhupriya

**Abstract** Due to the overwhelming adoption of cloud computing, the dire need to secure one's data is essential for a trustworthy cloud service provider. The key concerns of cloud security include unauthorized access, insecure interfaces, external sharing of data, malicious insiders, and denial-of-service attacks. Each business organization applies strong security mechanisms to protect sensitive data in the cloud. This includes the adoption of private cloud to prevent critical data from being accessible by outsiders.

**Keywords** Cloud computing · Security · Private cloud · Malicious · Sensitive

## 1 Introduction

Cloud computing is the technique of using services, applications, and data across the Internet. It comes in three different service models and four different deployment models. They are software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). The deployment models are public, private, community, and hybrid cloud.

### 1.1 Private Cloud Basics

**For protection, manageability, privacy, audit, compliance, and governance, private cloud yields a single point of control [1].** It can be hosted on-premises or remotely using company-owned hardware. For better results, a combined cloud model can be used [2]. Non-essential applications can be in the public cloud while maintaining applications that are important are hosted in-house [3].

---

D. I. George Amalarethinam (✉) · J. Madhupriya  
Department of Computer Science, Jamal Mohamed College (Autonomous), Affiliated to  
Bharathidasan University, Tiruchirappalli, Tamil Nadu 620020, India

## 1.2 Hybrid Cloud Security

Hybrid cloud security refers to the safeguarding of data, software, and infrastructure in an IT architecture that includes workload portability, orchestration, and management through multiple IT environments, including at least one public or private cloud. Better support for a remote workforce, lower costs, increased scalability and power, increased agility and innovation, business continuity, and enhanced protection and risk management are just a few of the benefits. The components of hybrid cloud security include physical, technical, and administrative layers. While physical security deals with security of vendors, technical security describes encryption and VPN. The administrative security is about rules and procedures that are clearly stated, such as processes for risk mitigation, plans for disaster management and policies on data security, and employee education [4].

For maintaining data confidentiality and integrity, cryptographic methods are to be applied. This includes symmetric key cryptography or public key cryptography; three basic goals are addressed through cryptographic techniques which are (i) authentication, (ii) confidentiality, and (iii) integrity. Integrity is a quality assuring the data does not alter between the source and destination [5]. Hybrid cloud, that combines public and private clouds, is seen as an option for storing healthcare data and includes a lot of private and confidential information [6].

The paper is organized as follows: Sect. 2 deals with Literature survey, Sect. 3 discusses the Problem Definition while Sect. 4 enumerates about the features of the SDE cipher. Section 5 lists the simulation results, Sect. 6 does the performance analysis, and Sect. 7 describes the concluding remarks.

## 2 Related Work

Madhiraju and Praveen Kumar [7] discussed about enterprise using a hybrid cloud to host their e-commerce Web site in a private cloud that is more secure and versatile, while their inventory site is hosted in the public cloud.

Varghese and Buyya et al. [8] elaborated that according to estimates, 63% of cloud users have adopted a hybrid cloud strategy, with use-cases reported in the healthcare and energy sectors. When using a public cloud, the points to be noted are bandwidth, latency, and network topologies.

Aryotejo et al. [9] conducted an investigation. The aim of this project is to evolve, check, and install an HCDM as a cloud infrastructure service. Here, metal as a service (MAAS) is used.

The open-source software own Gopala Krishnan and Uma Maheswari [10] introduced cloud. This enables file integration and sharing in a secure manner. It used Web Distributed Authoring and Versioning (WebDAV) to access files.

Liu et al. [11] took digital twin from concept to practice, introducing it as a precision simulation technology in industry. The paper focused on cloud healthcare

system design and digital twin health care (CloudDTH). This is a cloud-based, novel, generalized, and extensible framework for monitoring, diagnosing, and predicting aspects of an individual's health.

Vaishnnave et al. [12] in their paper elaborated about the four deployment models—public cloud, private cloud, community cloud, and hybrid cloud. The paper emphasized the advantages of hybrid cloud such as security, reliability, business coherence, management, and it narrated various techniques in cloud computing for security.

Hemdanet and Manjaiah [13] described a digital forensic approach to cybercrime detection in a private cloud setting. Using VMware's virtualization technology, an experimental setting was created. According to the results of the experiment, the proposed approach will support digital investigators and specialists in acquiring and extracting digital evidence from cloud storage system.

Sundarakani et al. [14] proposed a hybrid supply chain cloud approach, which combines infrastructure, tools, and platform configurations. This paper introduced a hybrid cloud to help with supply chain network integration.

Hamadah and Aqel [15] explored disaster recovery. It is a division of information technology that focuses on data security in the event of a disaster. It is a collection of procedures and tools for returning an object to normalcy following a catastrophic event. A recovery time target and a recovery point goal were the subject of the report.

Kanthan and Mohan [16] discussed in their white paper on the power of hybrid cloud strategy specified that security continues to be the top concern among organizations using public cloud, reported by over two-thirds (67%) of enterprises surveyed in this study. Hybrid cloud solutions resolve security, risk mitigation, and trust issues by allowing organizations to deploy workloads on its most suitable platform, based on compliance requirements.

Ghani et al. [17] elaborated about cloud storage issues, data security issues, and data management issues in detail. The authors in their paper explained about future cloud storage opportunities.

### 3 Problem Definition

Cloud computing, as an advanced technology, has posed various problems in information management. A user opting for hybrid cloud will prefer strong security mechanisms to protect his data available in both public and private clouds. Majority of the enterprises are moving toward hybrid cloud for potential benefits such as scalability, flexibility, cost efficiency, and security. **The private cloud in hybrid cloud ensures high degree of security.** This paper elucidates the symmetric data encryption (SDE) cipher for private cloud data protection in hybrid cloud. The proposed method is **SDE cipher—symmetric data encryption cipher, which uses symmetric key in the algorithm.** It encrypts data when it is stored.

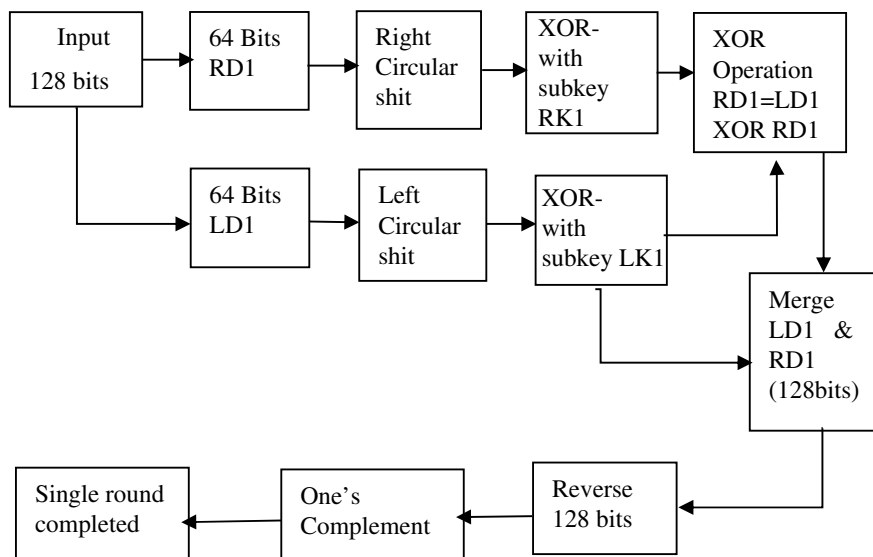
## 4 Features of SDE Cipher and Encryption Procedure

SDE cipher is a 128 bits block cipher. It uses two keys. The first key is the random number that denotes the number of rounds; the second key is the main key, from where three subkeys are generated for each round.

### 4.1 Encryption Procedure of SDE Cipher

- The users' data is the plaintext to upload.
- The plaintext is converted into binaries.
- Divide the total length of plaintext bits into 128 bits blocks.
- The SDE cipher takes the 128 bits binaries as input for encryption.
- If the plaintext has more than one 128 bits blocks, then every 128 bits are encrypted using the same subkeys.
- SDE cipher generates a round number key, which denotes the number of rounds.

### 4.2 Block Diagram of SDE



**Table 1** Comparison of proposed and existing techniques with respect to encryption time

Size	Encryption techniques		
	DES	3DES	SDE cipher
	Milliseconds		
1 MB	2843	7992	2132
2 MB	5690	15,982	4201
3 MB	8399	23,993	6313
4 MB	11,232	32,003	8421
5 MB	14,002	40,121	10,499

4.3 Sample Experiment with SDE Cipher Algorithm  
Encryption Procedure

The generated key from the cloud is Ω π %34^p#9ha7&\$t!

Plaintext: The hybrid cloud security.  
Ciphertext generated: z@\$;g^&#klw0a + y.  
Decryption procedure: The key same subkeys are used in reverse order for decryption.  
Ciphertext given is z@\$;g^&#klw0a + y.  
The plaintext: The hybrid cloud security.

5 Encryption and Decryption Time Comparison

**Simulation results:** The proposed encryption techniques are tested with different sizes of data and compared with existing algorithms. The proposed and current techniques take different amounts of time to encrypt data, as shown in Table 1.

The encryption time comparison of proposed and current encryption techniques is shown in Fig. 1. It demonstrates that the proposed SDE cipher encrypts data in the shortest time possible as compared to current techniques (Table 2).

The comparison of decryption times is shown in Fig. 2. The proposed SDE cipher took the least amount of time to decrypt the data, according to the results.

6 Analysis of SDE Cipher

A security-analyzing tool called ABC Hackman is used to quantify and evaluate security. The tool targets encrypted data in storage and attempts to recover original text contained in the ciphertext. The encryption algorithm’s security levels are defined in Table 3 and Fig. 3. The proposed SDE cipher achieves the highest level of security.

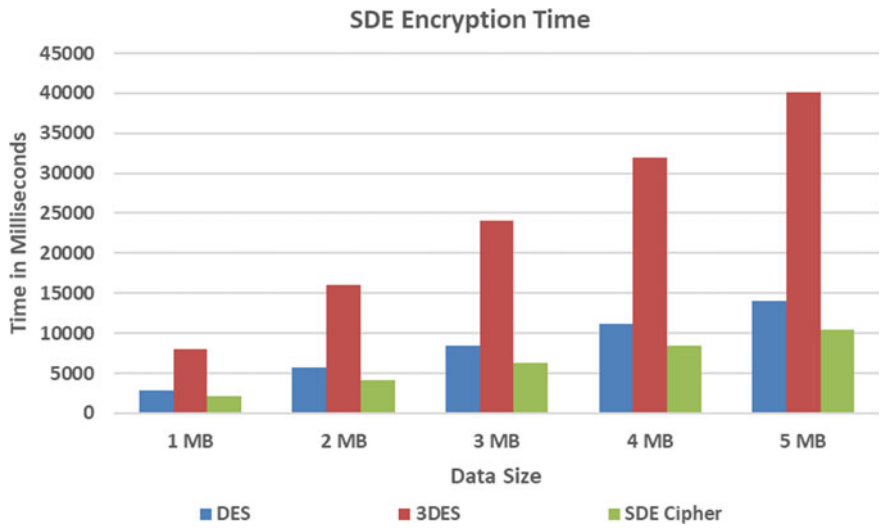


Fig. 1 SDE encryption time comparison

Table 2 Comparison of proposed and existing techniques with respect to decryption time

Size	Decryption techniques		
	DES	3DES	SDE cipher
	Milliseconds		
1 MB	2732	7992	2023
2 MB	5580	15,872	4101
3 MB	8290	23,883	6210
4 MB	11,140	31,911	8314
5 MB	13,911	40,012	10,390

As shown in the table and graph below, the percentage of security is calculated and compared.

7 Conclusion

The SDE cipher demonstrates that security in the private cloud is significantly improved. The private cloud is exclusively used for storing sensitive or critical data in the cloud. As such, it is important to implement stringent security measures. This SDE cryptographic algorithm is intended to improve security in private cloud environments. This methodology will be ideal for ensuring security in private clouds.



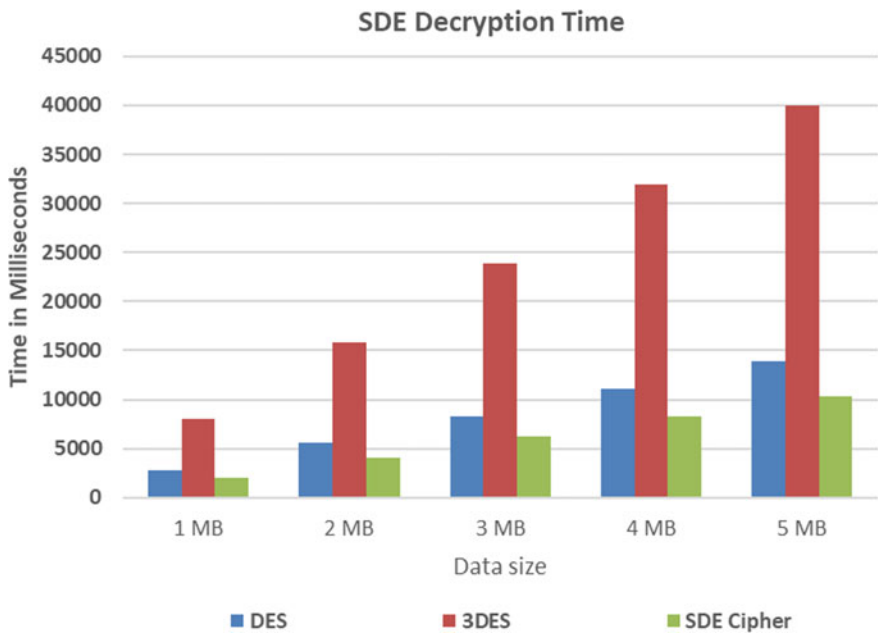
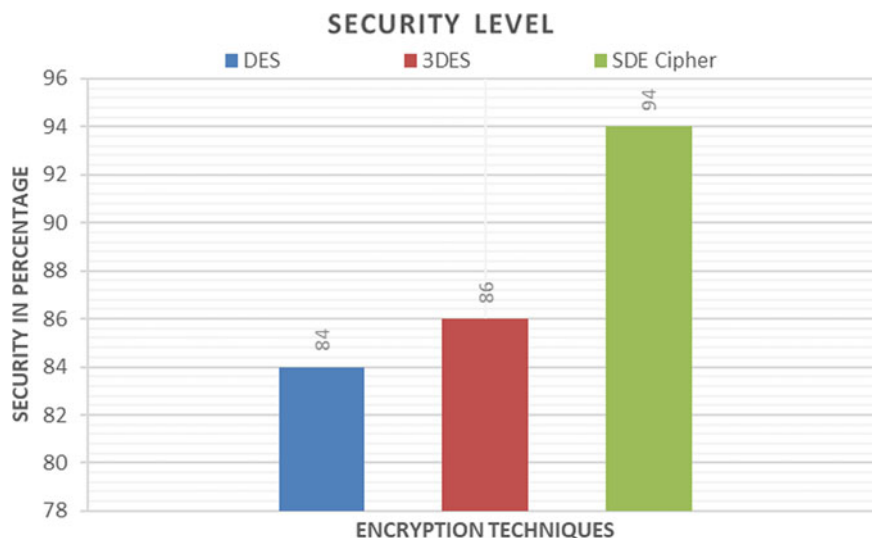


Fig. 2 SDE decryption time comparison

Table 3 Proposed and existing encryption techniques’ security levels

Security algorithms	Security level (%)
DES	84
3DES	86
SDE cipher	94



**Fig. 3** SDE security level comparison

## References

1. Kulkarni G, Patil N (2012) Private cloud secure computing. *International Journal of Soft Computing and Engineering (IJSCE)* 2(1). ISSN: 2231-2307
2. Dr. George Amalarethinam DI, Madhu Priya J (2020) MAGcipher: an enhanced cryptography encryption for securing the data in hybrid cloud. *Int J Adv Sci Technol* 29(5s):585–590
3. Owopetu OO (2013) Private cloud implementation and security using EUCALYPTUS and XEN frameworks. Bachelor's Thesis (UAS) Degree Program in Information Technology Internet Technology
4. Melnick J (2020) Hybrid cloud security. *Netwrix blog*
5. Josh RK, Kim J, Krumpholz A, Ranjan R, Shyamsundar RK, Nepal S (2019) Crowdsensing privacy in smart city applications. *Science Direct*
6. Omar AAI, Bhuiyan MZA, Basu A, Kiyomoto S, Rahman MS (2019) Privacy-friendly platform for healthcare data in cloud based on block chain environment. *Future Gen Comput Syst* 95:511–521
7. Priyanka Madhiraju P, Praveen Kumar M (2017) A research on simple way to a private cloud and its uses. *Int Res J Eng Technol (IRJET)* 04(01). e-ISSN: 2395-0056, p-ISSN: 2395-0072
8. Varghe B, Rajkumar Buyya R (2017) Next generation cloud computing: new trends and research directions. *Future Gen Comput Syst*. Elsevier Press, Amsterdam, The Netherlands. ISSN: 0167-739X
9. Aryotejo G, Kristiyanto DY, Mufadhol (2018) Hybrid cloud: bridging of private and Public cloud computing. *IOP Publishing, IOP Conf Series: Journal of Physics: Conf Series* 1025(2018):012091
10. Prakash Gopalakrishnan B, Uma Maheswari B (2019) Research on enterprise public and private cloud Service. *Int J Innov Technol Explor Eng (IJITEE)* 8(6S4). ISSN: 2278-3075
11. Liu Y, Zhang L, Yang Y, Zhou L, Ren L, Wang F, Liu R, Pang Z, Jamal Deen M (2019) A novel cloud-based framework for the elderly healthcare services using digital twin. *IEEE Access*
12. Vaishnnave MP, Suganya Devi K, Srinivasan P (2019) A survey on cloud computing and hybrid cloud. *Int J Appl Eng Res* 2:429–434. ISSN 0973-4562

13. Hemdan EED, Manjaiah DH (2018) Digital forensic approach for investigation of cybercrimes in private cloud environment. Springer
14. Sundarakani B, Kamran R, Maheshwari P, Jain V (2019) Designing a hybrid cloud for a supply chain network of Industry 4.0: a theoretical framework. Benchmarking, An International Journal
15. Hamadah S, Drah Agel D (2019) A proposed virtual private Cloud-Based Disaster Recovery Strategy. IEEE
16. Kanthan C, Deepak Mohan D (2019) White paper The Power of the Hybrid Cloud Strategy: addressing the breadth of enterprise Workload Requirements. Sponsored by: Dell EMC, Intel
17. Ghani A, Badshah A, Jan SU, Alshdadi AA, Daud A (2020) Department of computer science and software engineering, International Islamic University Islamabad, 44000, Pakistan. Issues and challenges in cloud storage architecture: a survey. Researchpedia J Comput 1(1):50–65

# Reducing Labelled Data Requirement for Pneumonia Segmentation Using Image Augmentations



Jitesh Seth, Rohit Lokwani, Viraj Kulkarni, Aniruddha Pant,  
and Amit Kharat

**Abstract** Deep learning semantic segmentation algorithms can localise abnormalities or opacities from chest radiographs. However, the task of collecting and annotating training data is expensive and requires expertise which remains a bottleneck for algorithm performance. We investigate the effect of image augmentations on reducing the requirement of labelled data in the semantic segmentation of chest X-rays for pneumonia detection. We train fully convolutional network models on subsets of different sizes from the total training data. We apply a different image augmentation while training each model and compare it to the baseline trained on the entire dataset without augmentations. We find that rotate and mixup are the best augmentations amongst rotate, mixup, translate, gamma and horizontal flip, wherein they reduce the labelled data requirement by 70% while performing comparably to the baseline in terms of AUC and mean IoU in our experiments.

**Keywords** Semantic segmentation · Augmentation · Chest X-rays · Medical image analysis · Deep learning

## 1 Introduction

The progress in computer vision has had a substantial impact on radiology [6]. Deep learning approaches such as convolutional neural networks (CNNs) have shown great success in the classification and segmentation of radiographs like chest X-rays and CT scans [4, 13]. Research on pathology identification in chest X-rays (CXRs) has mainly focussed on classification, where models predict a class label from a broad set of pathologies. Such an approach does not directly inform us of the regions in

---

J. Seth (✉)

Indian Institute of Science Education and Research, Pune, India

e-mail: [jitesh.seth@deeptek.ai](mailto:jitesh.seth@deeptek.ai)

J. Seth · R. Lokwani · V. Kulkarni · A. Pant · A. Kharat

DeepTek Inc, Pune, India

the CXRs responsible for the class label. It requires an additional step of plotting saliency maps or gradient class activation maps to interpret how the CNN is making decisions [17].

Segmentation and object detection algorithms provide an advantage over standard CNNs because they can predict the regions of interest. Architectures such as U-Net [20] and RetinaNet [14] localise the regions responsible for a specific pathology. Besides, semantic segmentation models are more useful than classification models in two ways: they require less training data since they have pixel-level labels for every image, and they can assist radiologists in their work by localising the abnormalities [9].

However, the cost and effort to label the data are often a significant constraint on training models that perform well in practice [18]. There have been efforts on both fronts—to increase the amount of labelled data available and to create techniques that can learn better from small amounts of labelled data. The former include large databases such as CheXpert [10] and ChestX-ray8 [26], whereas the latter consists of image augmentations, semi-supervised learning, special architectures such as U-Nets [20], GANs [8] and others.

Image augmentations are a vital data processing technique to improve the performance of machine learning models. They can make the CNN indifferent to naturally present variations in the data, such as position, scale or different radiography equipment. Augmentations can be of varied types. Rotate, scale and flip and similar augmentations are called geometrical augmentations. Photometric augmentations transform the colour space of the images. More complex transformations include elastic deformation or mixing multiple images. However, inflating the dataset with numerous augmentations would add to the training time and compute requirements without necessarily increasing performance. A drawback of augmentations is that they may cause overfitting by making the CNN invariant to some features but highly tailored to the training data in others [23].

The problem of augmentations on the performance of segmentation models in the medical domain is not sufficiently addressed in research. Knowledge of specific augmentations which reduce the labelled data requirements will help researchers and data scientists fine-tune their models faster and better. Moreover, data augmentation studies investigate the increase in model performance rather than the decrease in the labelled training data requirements. Our paper specifically addresses the reduction in the training data requirement using image augmentations.

In this paper, we implement five different augmentations on the training of CNN models on chest X-rays. We propose three criteria for identifying augmentations that reduce labelled data requirement. First, the model should perform comparably to the baseline on a subset of the data. Second, the models with augmentation trained on partial data should perform better than models without any augmentations trained on the same data. Third, the model should satisfy the criteria above for multiple test sets. We validated our results with an in-sample and out-of-sample test set.

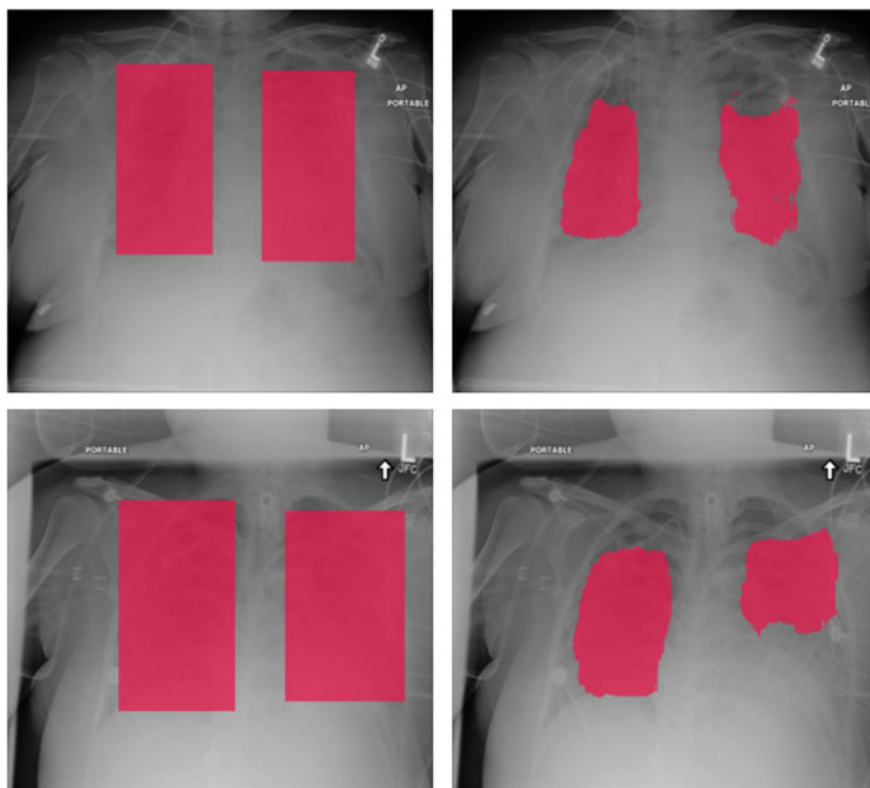
## 2 Related Work

This section covers the recent advancements in computer-aided diagnosis (CADx) using deep learning, image augmentations and semantic segmentation in the field of CXRs. CADx for chest radiography started around since 1960s [7], but deep learning has transformed and dominated the field in a few years. Van Ginneken et al. [6] has comprehensively summarised the evolution of CADx from rule-based and machine learning approaches to deep learning ones. Kermany et al. [12] showed the generalisability of deep neural networks by using the same neural network to classify retinal optical coherence tomography images and pediatric pneumonia classification in CXRs. The latter model achieved an accuracy of 92.8%, and the area under the ROC curve was 96.8%.

Recent studies have proven that image augmentations improve the machine learning model performance. Sirazitdinov et al. [24] showed the effect of augmentations on the classification of chest radiographs. They concluded that a combination of increasing brightness, random rotation and horizontal flips led to the best performance on the ChestX-Ray14 dataset, with an AUC ROC of 0.808 (compared to 0.785 without any augmentations). However, they do not quantify the extent of each augmentation, thus decreasing its reproducibility. After the invention of mixup [28], Eaton-Rosen et al. [5] applied it to a dataset of MRI images. They provide a graphical overview of mixup compared to other augmentations and a baseline for a large (199 images) and a small (10 images) dataset.

Souza et al. [25] created an automatic method for segmentation and reconstruction of lungs, which can take into account lung opacities from pneumonia or tuberculosis, reconstruct the lung boundaries and finally, segment the lungs. They used the segmented lungs for a classification model, which achieved an accuracy of 96.97%, an average Dice coefficient of 0.94 on the Montgomery County's tuberculosis control dataset [11]. Selvan et al. [21] tackled the same problem by treating high opacity regions as missing data and using a variational auto-encoder for data imputation. They achieved an accuracy of 88.15% and a Dice coefficient of 0.8503 on a curated CXR dataset. Thus, segmentation was mainly used to demarcate the lungs for use in classification models. However, we can also use semantic segmentation algorithms to demarcate lung opacities. This opportunity became plausible after the publication of the Radiology Society of North America's (RSNA) pneumonia dataset [22], also used in this study.

Wu et al. [27] took the concept of lung segmentation and opacity detection one step forward. They segmented both the lungs, divided them into three zones each and predicted the presence of pneumonia in each zone using the patient's radiology report. Thus, they created an object detection dataset from radiology reports. Using this dataset, they trained a RetinaNet model and tested it on the RSNA dataset. The model had a mean IoU of 0.29 per pneumonia positive image. Hurt et al. [9] have shown that semantic segmentation of CXRs can be used as a probability map to interpret the radiographs. Their segmentation model on the RSNA dataset showed a dice coefficient of 0.603, and the classification had an AUC of 0.854.



**Fig. 1** Two examples of CXR with the ground truth mask (left) and predicted mask by the baseline model (right)

An essential aspect of deep learning in CADx is the usability of the models—we do not desire clever models that might end up being clinically irrelevant [15]. For example, both in Pan and Cadrin-Chênevert’s, and Cheng’s model of the RSNA dataset [16], they systematically decreased the predicted bounding boxes by 12–17%, which increased the performance for the particular test set, but there is no medically relevant reason to do the same in practical settings. Another practical constraint with Pan’s models was extensive ensembling, which requires the availability of high-end GPUs.

### 3 Data and Methodology

#### 3.1 Data

For this study, we use the publicly available pneumonia dataset jointly annotated by RSNA and the Society of Thoracic Radiology (STR) [22]. The dataset consists of DICOM images of chest X-rays (CXRs) having dimensions  $1024 \times 1024$ . Pneumonia positive images contain bounding box information of lung opacities. The dataset consists of about 30,000 frontal CXRs with bounding boxes around the lung opacities.

From this dataset, we select all the pneumonia positive CXRs ( $n = 6012$ ) and a subset of the negative CXRs ( $n = 8488$ ) and divide them into training ( $n = 10,000$ ), validation ( $n = 1500$ ) and test ( $n = 3000$ ) sets. Each CXR belongs to a unique individual. All three sets have the same prevalence of the positive class (41.4%).

We also test all trained models on an out-of-sample test set curated from Padchest [1] and four private hospitals and population screening programmes from India and Indonesia. This set contains 1125 pneumonia positive CXRs with the corresponding polygonal or rectangular annotations and 1875 pneumonia negative CXRs.

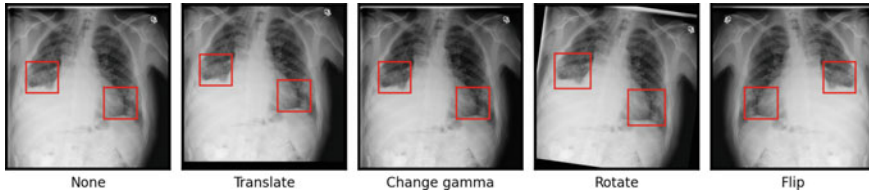
#### 3.2 Model Architecture and Evaluation

We use a U-net-like CNN with depthwise separable convolutions style connections implemented using the Keras library for this study [2]. We resize the input into an image of shape (512,512,3). The network has two parts, an encoder part and a decoder part. The encoder part uses residual connections, depthwise separable convolution and 2D convolution and max pooling. The decoder portion uses transpose convolution, 2D upsampling and 2D convolution. An example of the model prediction is given in Fig. 1.

We use binary cross-entropy as the loss function, and the evaluation metric is the mean intersection over union (IoU). We use Adam to optimise the loss function, and the learning rate is scheduled to decrease when plateauing for 5 epochs. We train each model until it has not shown improvement in validation loss for at least 5 epochs. We save the model weights with the lowest validation loss for evaluation.

After the model training, we evaluate it on the in-sample and out-of-sample test sets. These are compared to the ground truth masks to calculate the mean IoU and the loss. We use the segmentation results to classify the CXR as positive or negative. We calculate the area under the curve (AUC) of the receiver operating characteristics (ROC) based on this classification.





**Fig. 2** Original CXR with opacities labelled (left) and result of augmentations with updated bounding boxes

### 3.3 Augmentations

We have chosen five image augmentations for this study. These are random rotation between  $-10^\circ$  and  $10^\circ$ , changing the gamma between 0.75 and 1.25, translating the image randomly between 0-5% of its length in x- and y-directions, horizontal flips and mixup [28]. In mixup, two images and their masks (represented by  $x_1$  and  $x_2$ ) are combined using the formula:

$$x = x_1\lambda + x_2(1 - \lambda)$$

where  $\lambda \sim \beta(0.2, 0.2)$ . These augmentations have negligible computational cost. Mixup results in better-performing segmentation models according to recent studies [5].

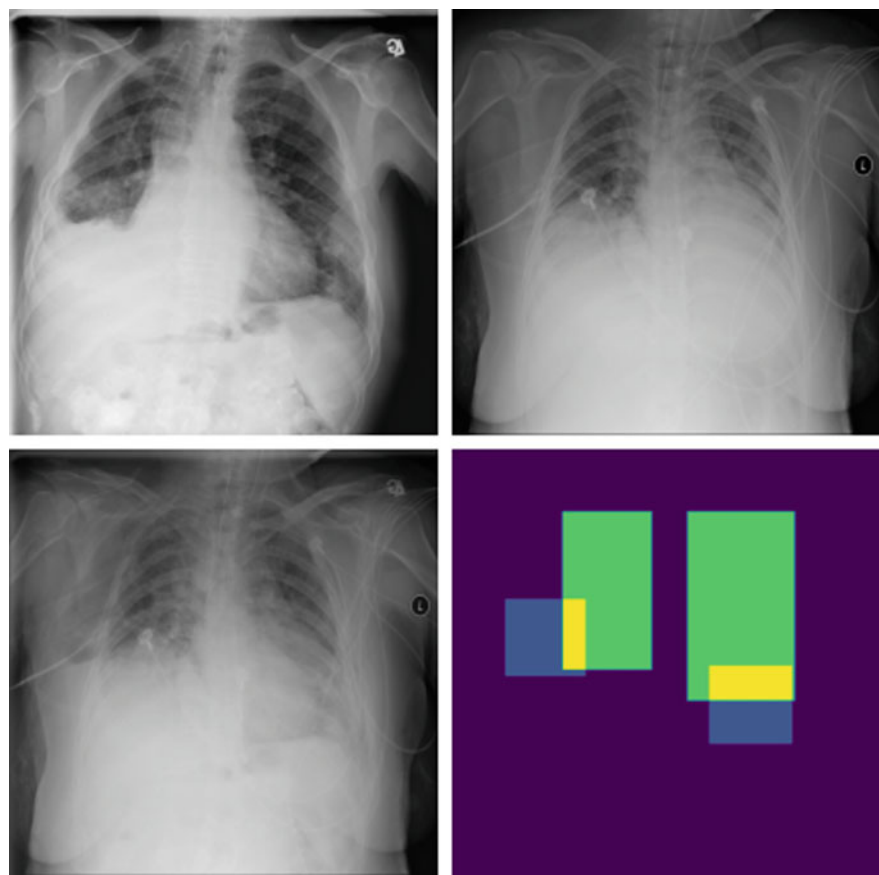
We train the baseline (i.e. no augmentations) on 100% of the training set. For each of the six conditions—five augmentations and one with no augmentation (hereafter referred to as “NoAug”), models were trained using 30, 50, 70 and 90% of the training set (Figs. 2 and 3).

### 3.4 Statistical Analysis

We compare each model with the baseline and the NoAug model on the same amount of data. We use the nonparametric DeLong test [3] to compare the AUC for the models’ classification performance. We set the significance level at 0.05. We use the pROC library in R [19] to perform the same.

## 4 Results

For comparison of the models’ segmentation performance, we plotted each model’s mean IoU in Fig. 4. We also compared the different models’ classification performances by using AUC, in Fig. 5.

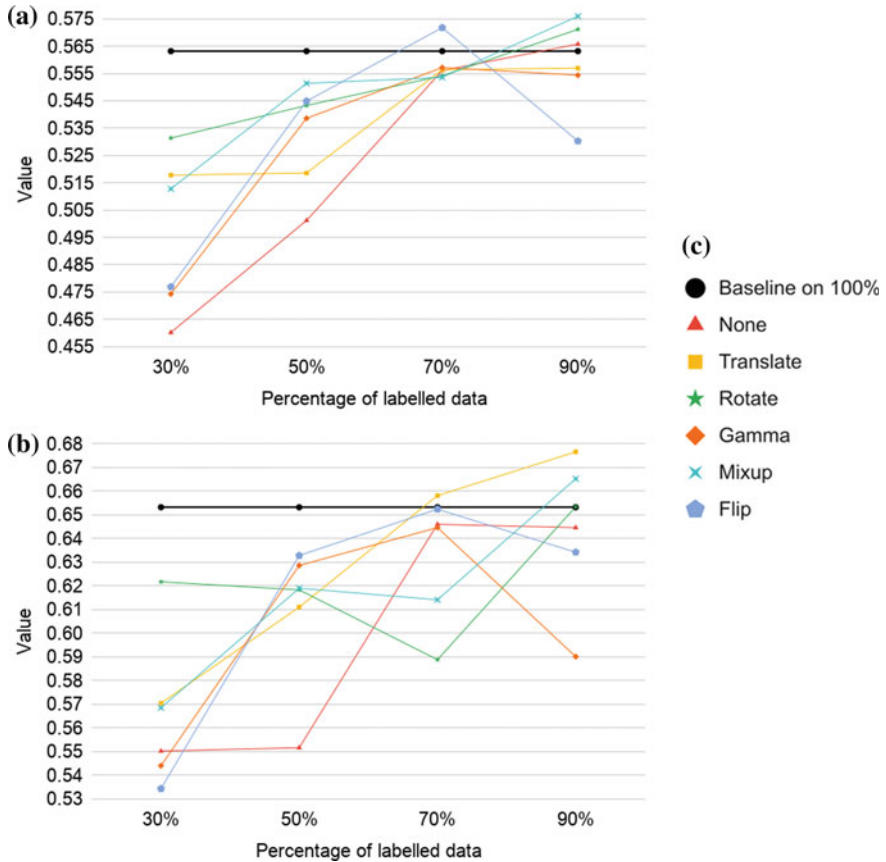


**Fig. 3** (Top) original CXRs for mixup; (bottom left) result of mixup augmentation and (bottom right) corresponding mask

**Table 1** p-values of DeLong test compared to the baseline for the external test set

Test set	Internal test set		External test set	
	30%	50%	30%	50%
Augmentation	30%	50%	30%	50%
None	$7.66 \times 10^{-12} / 1$	0.4826/0.7587	$2.20 \times 10^{-16} / 1$	$6.04 \times 10^{-8} / 1$
Translate	<b>0.2288</b> /0.1144	<b>0.176</b> /0.912	$2.34 \times 10^{-5} / 1$	<b>0.084</b> /0.958
Rotate	<b>0.1576</b> /0.0788	<b>0.6309</b> /0.3154	<b>0.0599</b> / <b>0.0300</b>	<b>0.2606</b> /0.8697
Gamma	0.0147/0.9927	1.86/ <b><math>9.29 \times 10^{-4}</math></b>	$1.11 \times 10^{-13} / 1$	$5.15 \times 10^{-4} / $ <b><math>2.58 \times 10^{-4}</math></b>
Mixup	<b>0.908</b> /0.454	<b>0.0546</b> / <b>0.0273</b>	$2.70 \times 10^{-8} / 1$	<b>0.4118</b> /0.2059
Flip	<b>0.7151</b> /0.3575	0.02505/0.9875	<b>0.2316</b> /0.8842	$6.984 \times 10^{-4} / 0.9997$

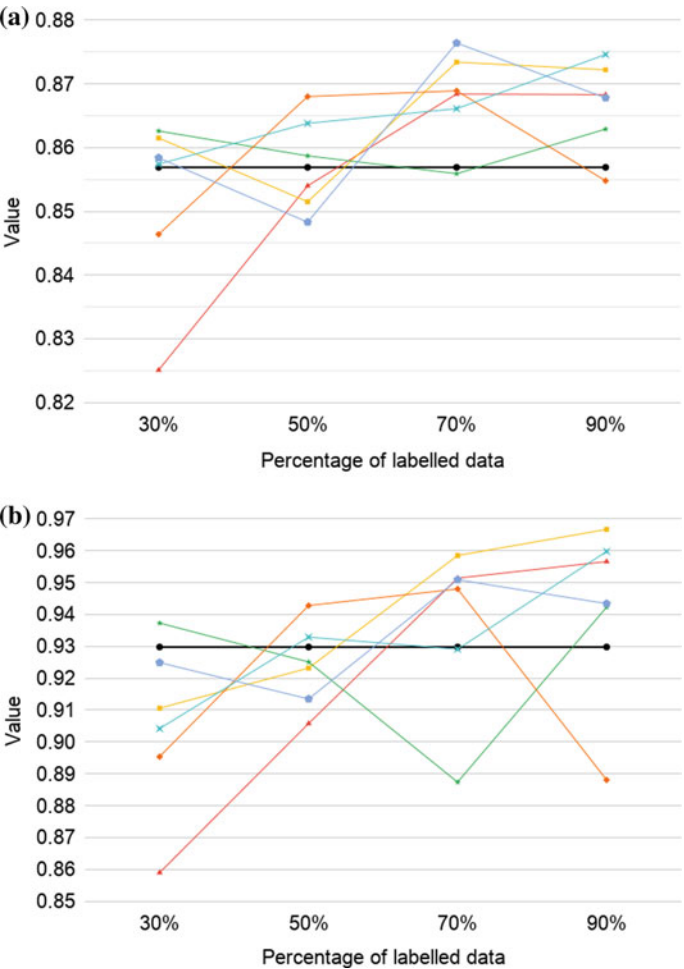
The first value corresponds to two-tailed DeLong test and the second corresponds to the one-tailed DeLong test. Bold indicates that the value satisfied our criteria, as explained in Sect. 4



**Fig. 4** Mean IoU on the **a** internal test set and **b** external test set for the various models trained. **c** The legend

It is essential to note the behaviour of the NoAug models. In both Figs. 4 and 5, we see that the performance of NoAug models is lower than baseline for 30% and 50% data. However, the NoAug models with 70% and 90% data perform as good as, or even better, than the baseline. Thus, it is irrelevant to study augmentations on 70% or more fraction of the data for this study. The p-values for the one-tailed DeLong test between AUC of NoAug and baseline are  $\sim 10^{-3}$  and  $\sim 10^{-4}$ , respectively, for the internal test set.

To check the first criterion, we did two DeLong hypothesis tests to see if the augmentation models' performance is comparable to the baseline. The first has the null hypothesis that the AUC of augmentation models is equal to that of the baseline trained on 100% data. In this case, a p-value larger than 0.05 would mean that we fail to reject the null hypothesis, and thus, we can say the two AUC are comparable. In the second test, the null hypothesis is that the AUC of the augmentation models is lesser than the baseline. Here a p-value of less than 0.05 would mean that we



**Fig. 5** AUC ROC of the **a** internal and **b** external test set for the various models

can reject the null hypothesis, and the AUC of augmentation models is significantly larger than that of the baseline. We find that except for gamma with 30% data and flip with 50% data, all models trained on 30% and 50% data pass either of the two tests on the internal test set. However, for the external test set, only rotate and flip with 30% data, and translate, rotate, gamma and mixup with 50% data pass the hypothesis tests (see Table 1 for the p-values).

We checked the second criterion by performing a one-tailed DeLong Test on the AUC of the augmentation against NoAug models trained on the same data on both test sets. We find that for the external test set, all augmentations with 30% and 50% data perform significantly better than NoAug ( $p < 10^{-2}$ ). For the internal test set, all augmentations trained on 30% data performed better than NoAug with 30% data. However, on 50% data, only the gamma and mixup performed better than NoAug.

**Table 2** AUC ROC of the different augmentations calculated on 30 and 50% data for both internal and external test sets

Test set	Internal test set		External test set	
Augmentation	30%	50%	30%	50%
Baseline on 100%	0.8569	0.8569	0.9298	0.9298
None	0.8251	0.854	0.859	0.9058
Translate	0.8615	0.8515	0.9106	0.9232
Rotate	0.8626	0.8587	0.9373	0.9251
Gamma	0.8464	0.868	0.8954	0.9428
Mixup	0.8574	0.8638	0.9042	0.9329
Flip	0.8584	0.8483	0.9249	0.9135

We propose that good augmentations should satisfy both the criteria above for both the test sets. We found that rotate and flip with 30% data and gamma and mixup with 50% data accomplished this (Table 2).

On the other hand, the mean IoU plot informs us of the best data augmentations at pixel level. As we can see in the internal test set (Fig. 4a), rotate and mixup perform quite good at 30%, whereas mixup and flip perform better at 50%. For the external test set (Fig. 4b), we see that rotate and mixup at 30% and flip and gamma at 50% performing better than NoAug and almost as good as the baseline.

Therefore, we find that rotate and mixup are the best augmentations for semantic segmentation on our dataset. These augmentations are capable of reducing the labelled data requirements by even 70%.

## 5 Conclusion

Users of deep learning algorithms often overlook image augmentations as an extra step for a minor boost in performance. Our study has shown that augmentations are capable of reducing the amount of labelled training data required. To the best of our knowledge, this is the first study addressing augmentations in this manner.

There is an assumption that augmentations are most useful when the training and test data are from the same distribution [23]. Using two test sets from different countries (internal test set sourced from the USA and the out-of-sample test set from India) with different radiology protocols and image qualities, we still achieved good results on both. Thus, we have shown that the assumption need not hold.

While this study looked at individual augmentations, there is still scope for improving the model performance and reducing the labelled data requirement further by combining multiple augmentations. This study contributes to the field of CADx by enabling practitioners/researchers in choosing the right augmentations for building segmentation models when there is limited data of varied quality.

## References

1. Bustos A, Pertusa A, Salinas JM, de la Iglesia-Vayá M (2020) PadChest: a large chest x-ray image dataset with multi-label annotated reports. *Med Image Anal* 66:101,797. <https://doi.org/10.1016/j.media.2020.101797>. <http://arxiv.org/abs/1901.07441>. ArXiv: 1901.07441
2. Chollet F (2020) Keras documentation: image segmentation with a U-Net-like architecture. Keras website. [https://keras.io/examples/vision/oxford\\_pets\\_image\\_segmentation/](https://keras.io/examples/vision/oxford_pets_image_segmentation/). [https://keras.io/examples/vision/oxford\\_pets\\_image\\_segmentation/](https://keras.io/examples/vision/oxford_pets_image_segmentation/). Accessed: 28-Dec-2020
3. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 837–845
4. Dunnmon JA, Yi D, Langlotz CP, Ré C, Rubin DL, Lungren MP (2018) Assessment of convolutional neural networks for automated classification of chest radiographs. *Radiology* 290(2):537–544
5. Eaton-Rosen Z, Bragman F, Ourselin S, Cardoso MJ (2018) Improving data augmentation for medical image segmentation (2018). <https://openreview.net/forum?id=rkBBCjjiG>
6. van Ginneken B (2017) Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. *Radiolog Phys Technol* 10(1):23–32. <https://doi.org/10.1007/s12194-017-0394-5>
7. Ginneken BV, Romeny BMTH, Viergever MA (2001) Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans Med Imag* 20(12):1228–1241. <https://doi.org/10.1109/42.974918>
8. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial networks. arXiv:1406.2661 [cs, stat] (2014). <http://arxiv.org/abs/1406.2661>. ArXiv: 1406.2661
9. Hurt B, Yen A, Kligerman S, Hsiao A (2020) Augmenting interpretation of chest radiographs with deep learning probability maps. *J Thor Imag* 35(5):285–293
10. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, Marklund H, Haghighi B, Ball R, Shpanskaya K, Seekins J, Mong DA, Halabi SS, Sandberg JK, Jones R, Larson DB, Langlotz CP, Patel BN, Lungren MP, Ng AY (2019) CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. arXiv:1901.07031 [cs, eess] (2019). <http://arxiv.org/abs/1901.07031>. ArXiv: 1901.07031
11. Jaeger S, Candemir S, Antani S, Wang YXJ, Lu PX, Thoma G (2014) Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imag Med Surg*. 4(6):475–477
12. Kermay DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang ED, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172(5):1122–1131.e9
13. Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* 284(2):574–582
14. Lin TY, Goyal P, Girshick R, He K, Dollár P (2018) Focal loss for dense object detection. arXiv:1708.02002 [cs] (2018). <http://arxiv.org/abs/1708.02002>. ArXiv: 1708.02002
15. Lundervold AS, Lundervold A (2019) An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik* 29(2):102–127
16. Pan I, Cadrin-Chênevert A, Cheng PM (2019) Tackling the radiological society of North America Pneumonia detection challenge. *Am J Roentgenol* 213(3):568–574
17. Pasa F, Golkov V, Pfeiffer F, Cremers D, Pfeiffer D (2019) Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization. *Sci Rep* 9(1):6268. <https://doi.org/10.1038/s41598-019-42557-4>

18. Prevedello LM, Halabi SS, Shih G, Wu CC, Kohli MD, Chokshi FH, Erickson BJ, Kalpathy-Cramer J, Andriole KP, Flanders AE (2019) Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. *Radiol.: Artif. Intell.* 1(1), e180,031 (2019). 10.1148/ryai.2019180031. <https://pubs.rsna.org/doi/full/10.1148/ryai.2019180031>
19. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf* 12:77
20. Ronneberger O, Fischer P, Brox T (2019) U-Net: convolutional networks for biomedical image segmentation. [arXiv:1505.04597](https://arxiv.org/abs/1505.04597) [cs] (2015). <http://arxiv.org/abs/1505.04597>. [ArXiv: 1505.04597](https://arxiv.org/abs/1505.04597)
21. Selvan R, Dam EB, Detlefsen NS, Rischel S, Sheng K, Nielsen M, Pai A (2020) Lung segmentation from chest X-rays using variational data imputation. [arXiv:2005.10052](https://arxiv.org/abs/2005.10052) [cs, eess, stat] (2020). <http://arxiv.org/abs/2005.10052>. [ArXiv: 2005.10052](https://arxiv.org/abs/2005.10052)
22. Shih G, Wu CC, Halabi SS, Kohli MD, Prevedello LM, Cook TS, Sharma A, Amorosa JK, Arteaga V, Galperin-Aizenberg M, Gill RR, Godoy MC, Hobbs S, Jeudy J, Laroia A, Shah PN, Vummi D, Yaddanapudi K, Stein A (2019) Augmenting the national institutes of health chest radiograph dataset with expert annotations of possible pneumonia. *Radiol Artif Intell* 1(1):e180,041. <https://doi.org/10.1148/ryai.2019180041>. <https://pubs.rsna.org/doi/abs/10.1148/ryai.2019180041>
23. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. *J Big Data* 6(60). <https://doi.org/10.1186/s40537-019-0197-0>
24. Sirazitdinov I, Kholiavchenko M, Kuleev R, Ibragimov B (2019) Data augmentation for chest pathologies classification. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), pp 1216–1219. <https://doi.org/10.1109/ISBI.2019.8759573>. ISSN: 1945-8452
25. Souza JC, Diniz JOB, Ferreira JL, da Silva GLF, Silva AC, de Paiva AC (2019) An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Comp Methods Prog Biomed* 177:285–296
26. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) ChestX-Ray8: hospital-scale chest X-Ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), pp 3462–3471. <https://doi.org/10.1109/CVPR.2017.369>. ISSN: 1063-6919
27. Wu J, Gur Y, Karargyris A, Syed AB, Boyko O, Moradi M, Syeda-Mahmood T (2020) Automatic bounding box annotation of chest X-Ray data for localization of abnormalities. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI), pp 799–803 (2020). <https://doi.org/10.1109/ISBI45749.2020.9098482>. ISSN: 1945-8452
28. Zhang H, Cisse M, Dauphin YN, Lopez-Paz D (2018) mixup: beyond empirical risk minimization. [arXiv:1710.09412](https://arxiv.org/abs/1710.09412) [cs, stat] (2018). <http://arxiv.org/abs/1710.09412>. [ArXiv: 1710.09412](https://arxiv.org/abs/1710.09412)

# Feature Extraction for Automatic Grading of Students' Presentations



Jyoti G. Borade, Arvind W. Kiwelekar, and Laxman D. Netak

**Abstract** This paper presents an automated approach to assess the students' presentations. The approach aims to evaluate students' knowledge about various functionalities provided by presentation software. The method uses the *Python-pptx* library to extract description-specific features present in the presentations. Multiple machine learning techniques are used to grade presentations. Further, the approach uses a tree-based feature selection method to remove the least important features. A dimensionality reduction technique is used to improve the performance of the machine learning model.

**Keywords** Feature extraction · Automated grading · Feature selection and reduction · Assessment

## 1 Introduction

Students communicate design ideas through PowerPoint presentations. Hence, delivering presentations is an important academic activity. Presentations have both technical and non-technical aspects. Evaluators judge technical elements, including the expertise of a student in the field. Evaluators also judge non-technical aspects referred to as presentation quality.

However, evaluating presentations is a tedious task when the number of students is enormous. In such cases, evaluation depends on the evaluator's knowledge and experience. Examiners can use rubrics to overcome subjectivity in the assessment. But, many evaluators prefer to use their expertise to grade students' presentations instead of using pre-defined rubrics. Hence, it is a time-consuming mechanical activity and susceptible to inconsistencies and inaccuracies [1].

---

J. G. Borade (✉) · A. W. Kiwelekar · L. D. Netak  
Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University,  
Lonere, Raigad 402103, India  
e-mail: [awk@dbatu.ac.in](mailto:awk@dbatu.ac.in)

L. D. Netak  
e-mail: [ldnetak@dbatu.ac.in](mailto:ldnetak@dbatu.ac.in)



We present a data-driven approach to overcome these challenges. The approach aims to grade students' presentations based on presentation quality and not on technical content or subject matter included in presentations. Our approach is built on a data set of existing PowerPoint slides prepared by students and grade assigned by examiners based on presentation quality. It extracts the features that contribute to enhancing the quality of a presentation. We have used features like images, charts, graphs, tables, fonts, colors, hyperlinks, footers, etc., which captures various aspects enhancing the quality of presentations. It is supposed that human experts will grade presentations based on technical aspects. Our main objective is to evaluate the efforts put by the students in preparing PowerPoint presentations.

The structure of this paper is as follows. Feature extraction and selection techniques are discussed in Sect. 2. Section 3 presents the design of our system. Related work is described in Sect. 4. The paper concludes in Sect. 5.

## 2 Feature Engineering: Overview

This section briefly reviews various methods involved in identifying the relevant and optimal features from the existing data set. The machine learning models suffer from a phenomenon called *curse of dimensionality* when the number of features in the data set is enormous because it increases space and the time complexity of the machine learning model. Hence, selecting the most relevant features is crucial for getting the optimal performance of machine models. It includes mainly two processes called feature selection and feature reeducation methods.

**(A) Feature selection:** The activity of feature selection is performed to select a subset of the features. Feature selection methods remove redundant and irrelevant features. The commonly used feature selection techniques are (i) *filter methods*: It removes duplicate, correlated and redundant features by measuring information gain, conducting a Chi-square test and calculating variance threshold. (ii) *Wrapper methods*: It takes a subset of the features to train the model. The subset of features that gives the best performance is considered the optimal set of features. Wrapper methods provide better accuracy compared to filter methods but computationally more expensive. Some of the commonly used methods include forward, backward and bidirectional elimination methods. (iii) *Embedded methods*: It removes the drawback of filter and wrapper methods. These methods are faster like filter methods and accurate like wrapper methods. These are two kinds, namely tree-based methods and regularization (e.g., LASSO, ridge regression).

**(B) Feature reduction:** Feature selection is often mistaken as dimensionality reduction. Both methods reduce the number of features in the data set. But in dimensionality reduction, features are transformed into new combinations of features. Feature selection methods select or reject features without any change in them. A few of the commonly used dimensionality reduction techniques are discussed below: (i) *Principal component analysis (PCA)*: It transforms a set of correlated features into a set of uncorrelated features. It helps to transfer complex data set into a sim-

ple one. It finds a strong correlation between various features. To find the principal components, methods use operations such as standardizing the data set, computing the covariance matrix and calculating the eigenvectors and eigenvalues.

(ii) *Linear discriminant analysis (LDA)*: It is also used to reduce the number of features in the data set. It holds complete information about the features. It creates a new axis and projects the information on it to minimize variance and maximize the distance between the mean of two classes.

LDA is a supervised method of feature selection, whereas PCA is unsupervised. PCA has preferred when classwise samples are few, and LDA is preferred when multiclass classification having a large data set is to be performed.

### 3 System Architecture

Figure 1 shows the architecture of our system designed to carry out automatic grading of PowerPoint presentation. Solid lines show training process of the model. Dotted lines show prediction process of our model. A data set of twenty-five PowerPoint presentations has been collected. Two teachers have separately validated the correctness of extracted features. The PowerPoint presentation slides have been graded separately by two different evaluators. These features are not disclosed to them. With their permission, we have recorded grades in the data set as output label. We have implemented feature extraction program to extract features as shown in Table 1 using *Python – pptx* library. We have optimized features by removing unimportant features. To carry out this reduction process, we have used tree-based method, *ExtraTreeClassifier* which is an ensemble learning technique. Features’ dimensionality is reduced by using linear discriminant analysis. The machine learning models have been implemented using *scikit – learn* library provided by Python. Further, models are tested by passing unseen PowerPoint presentation to the learned model to get grade.

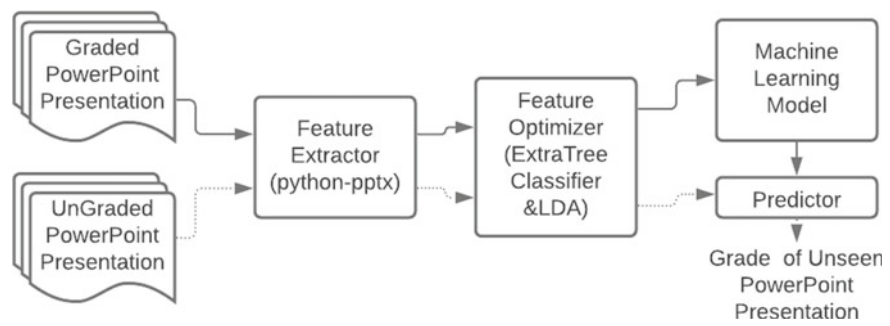


Fig. 1 System Architecture

**Table 1** List of the PowerPoint presentation features extracted using *Python – pptx* library

Feature Group	Features	Remark
Graphics	Image	This is a Boolean data checking whether the element is an image, diagram or a figure
	Chart	A Boolean data checking whether charts are used
	Table	A Boolean data indicating presence of a table
	Textbox	Checks whether text boxes are used or not
	Placeholder	A Boolean data checking presence of placeholders
Text appearance	Maximum font size	This is numeric variable returns maximum size of text font used in the paragraph
	Minimum font size	This is numeric variable returns minimum size of text font used in the paragraph
	Bold	A Boolean variable checks for bold formatting
	Underline	A variable checks for underline formatting
	Italic	This Boolean variable checks the use of italic text
	Number of font types	It returns numeric value indicating number of font types used in paragraph
	Number of font size	It returns numeric value indicating number of font size used in paragraph
	Number of font colors	This is numeric variable returns number of font colors used for text present in the paragraph
Footer Hyperlinks	Footer	This is Boolean indicates presence of footer
	Date	A Boolean data checking the date on the slide
	Hyperlink	A Boolean data checking for hyperlinks
	Number of hyperlinks	This is a numeric value indicates number of hyperlinks used in the presentation
	Slide number	This is a numeric variable returns slide number
	Number of slides	This is numeric variable returns number of slides present in the presentation

### 3.1 Feature Extraction Using Python-Pptx

The *Python – pptx* is a library supported by Python, and it is used for creating and manipulating PowerPoint presentations from Python programs. It has capabilities to support following features in the presentation: (i) (ii) Add new slides in the presentation, (iii) Populate text placeholders to create bullet slides, (iv) Add images to a slide at arbitrary position, (v) Add text box to a slide. manipulate text font size and make it bold, (vi) Add tables to a slide, (vii) Add auto shapes (e.g., polygons, flowchart, etc.) to a slide, (viii) Add and manipulate column, bar, line and pie charts and (ix) Access and change core document properties such as title and subject.

The placeholder is an important concept supported by PowerPoint slides. It is a shape into which different contents such as tables, charts, SmartArts can be placed. Some of the placeholder types are as follows:

1. **Title, Center Title, Subtitle, Body:** These groups of placeholders contain text.
2. **Content:** It allows insertion of a table, a chart, SmartArt, a picture, clip art or a media clip.
3. **Date, Footer, Slide Number:** Using them we can identify whether date, footer and slide number are present in presentations or not.
4. **Vertical Body, Vertical Object, Vertical Title:** They are useful for vertically oriented languages like Japanese.

We have collected the presentation slides prepared by students to deliver a course seminar. Students can prepare PowerPoint slides using any presentation software (e.g., *MSOffice*, *Python – pptx* library). No separate training is provided to students on how to prepare presentations. Their performance in the seminar is typically evaluated against the efforts put by a student to prepare PowerPoint slides, coverage of the topic selected by a student and communication skills of a student. Our approach checks the efforts put by a student to prepare PowerPoint presentation and not for the topic and communication skills. We have collected about twenty-five PowerPoint presentations. These slides are used to extract features. Various machine learning models are built to do automatic grading of presentations. Feature extraction and selection program as well as machine learning models have been made available on GitHub ([https://github.com/jyotiborade/feature\\_extraction\\_grading/](https://github.com/jyotiborade/feature_extraction_grading/)).

We have identified following features to evaluate the quality of presentations. These are as follows: (i) *Text appearance:* As shown in Table 1, total eight features have included in this group. These features mainly capture the information about the appearance of text on the slide. We presume that a diverse presentation in terms of text color, font size and font type. Usually, these features aim to attract the audience's attention. (ii) *Graphics:* This set of features captures the information about the usage of images, graphs, charts, etc., in presentation. A presentation that includes images, diagrams, graphs and charts usually conveys information effectively. There are five different features included in this category as listed in Table 1. (iii) *Footer and hyperlinks:* These features capture information about presence of footer, date and slide number. It indicates whether hyperlinks are provided to smoothly move

across the presentation. (iv) *Output Features*: We have graded presentation in four categories as *Excellent*, *Very Good*, *Good* and *Fair*. These values are assigned by a human evaluator.

3.2 Model Evaluation

Typically, the quantitative metrics such as *F1-score*, *Precision*, *Recall* and *Accuracy* are used to evaluate classifier models. The bar charts in Fig. 2 and Table 2 show accuracy of various machine learning classifiers. Decision tree (DT) and support vector machine (SVM) show 100% accuracy while predicting the class of presentations. In order to improve performance of multilayer perceptron (MLP) and Naive Bayes (NB) classifier, we have used embedded tree-based method for feature selection. It selects features which are more important in making an impact on the target feature. We have selected 19 features of high importance. Presence of placeholder, title, hlink and italic font is removed from our feature set. After applying tree-based feature

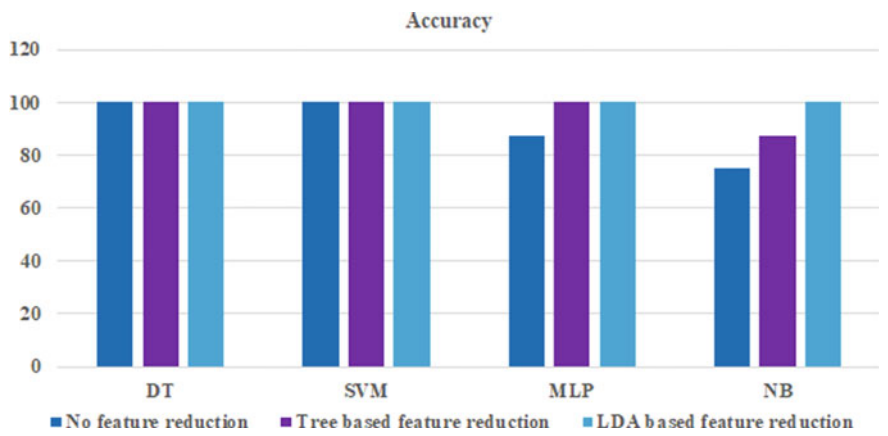


Fig. 2 Accuracy

Table 2 Accuracy

Machine learning model	Accuracy		
	No feature reduction	Tree-based feature reduction	LDA-based feature reduction
DT	100	100	100
SVM	100	100	100
MLP	87.50	100	100
NB	75	87.5	100

reduction for MLP, it got 100% accuracy. But for NB, performance is not improved. Hence, further we have applied LDA dimensionality reduction technique to get 100% accuracy for Naive Bayes classifier. Table 3 shows *Precision*, *Recall* and *F1-score* of various classifiers for different classes. We can conclude that precision, recall and F1-score are improved after applying feature selection and reduction techniques.

## 4 Related Work and Discussion

Developing computer-based automated techniques for assessment is referred by multiple terms such as automated grading of essays (AGE) [2] and computer Automated scoring [3]. Automatic assessment is done by using various machine learning models [4]. The performance of machine learning models depends on the quality and quantity of the features used. Researchers have extracted features using NLP techniques [5] and deep learning techniques [6]. Researchers have also developed automated scoring for many natural languages like Indonesian [7] are Chinese [8] are to name a few. These automated scoring techniques assess various skills such as essay writing, physician's patient management skills, dentistry assessment skills and architects registration process.

In this paper, we have presented an approach to assess the students' PowerPoint presentation skills. The main objective is to assess student's skill which they have used to prepare PowerPoint presentation. For this purpose, we have used features provided by *Python – pptx* library in Python. We have implemented feature extraction program to prepare the data set from the PowerPoint slides prepared by students. Then we have used multiple machine learning techniques to automate the task of grading. Performance of machine learning classifier is improved by using feature selection and dimensionality reduction techniques. It relies upon the description-specific presentation features to grade PowerPoint presentation.

## 5 Conclusion

We have implemented a technique in which PowerPoint presentation skill of students can be assessed using computer-based automated techniques. We have implemented feature extraction program using *Python – pptx* library to identify a set of useful features to determine the presentation quality. Using this feature extraction program, we have developed a small data set to enable the development of machine learning techniques. A data-driven approach to assess the presentation skill is demonstrated through various prototype classifiers. Unnecessary features are removed using tree-based feature selection method. Also, we have used linear discriminant analysis technique to do dimensionality reduction. After applying feature selection and dimensionality reduction techniques, we have achieved 100% accuracy for all the classifiers to predict the grade of PowerPoint presentation.



The performance of machine learning-based model depends on the set of input features used. Feature extraction is done to select features which contribute a lot to our output. We have to identify minimal number of relevant features to increase accuracy of our machine learning model. Natural language processing and deep neural networks-based approaches like auto-encoders can be developed for automatic feature extraction [9–11]. Many programming languages such as Java, JavaScript and Python provide libraries to process PowerPoint presentations created in *PPT* format. We have designed program for feature extraction using *Python* – *pptx* library of Python.

## References

1. Jyoti G, Borade LD (2020) A review, IHCI, Netak, Automated grading of essay
2. Hearst MA (2000) The debate on automated essay grading. *IEEE Intelligent Syst Appl* 15(5):22–37
3. Yang Y, Buckendahl CW, Juskiewicz PJ, Bhola DS (2002) A review of strategies for validating computer-automated scoring. *Appl Measurement Educ* 15(4):391–412
4. Ullmann TD (2019) Automated analysis of reflection in writing: validating machine learning approaches. *Int J Artif Intelligence Educ* 29(2):217–257
5. Aluizio HF, Hércules P, Edilson F, Jonathan N (2018) An approach to evaluate adherence to the theme and the argumentative structure of essays. *Proc Comput Sci* 126: 788–797
6. Surya K, Ekansh G, Kailasanathan, Nallakaruppan (2019) Deep learning for short answer scoring. *Int J Recent Technol Eng* 7: 1712–1715
7. Faisal R, Trisna R, Rohadi E, Siradjuddin, Indrazno, Rudy A, Awan S, Supriatna A (2018) Open Problems in Indonesian automatic essay scoring system. *Int J Eng Technol* 7: 156–160
8. Peng X, Ke D, Chen Z, Xu B (2010) Scoring automated Chinese essay, using vector space models. In: 4th International Universal Communication Symposium. Beijing, pp 149–153
9. Abdallah B, Abubakr H, Benjamin R, Daniel D, Mohanad A (2018) Implementation of a neural natural language understanding component for Arabic dialogue systems. *Proc Comput Sci* 142:222–229
10. Ramalingam VV, Pandian A, Prateek C, Himanshu N (2018) Automated essay grading using machine learning algorithm. *J Phys: Conference Series*. 1000. 012030. 2018
11. George N, Sijimol PJ, Varghese SM (2019) Grading descriptive answer scripts using deep learning. *Int J Innov Technol Exploring Eng (IJITEE)* ISSN: 2278-3075, vol 8, issue 5, March 2019



# Monitoring and Assessment of Water Bodies in Madurai Using Geo-Spatial Data



V. Vignaraj Ananth, R. Sarunisha, S. Aditya, M. Haripriya,  
and M. Bhuvaneshwari

**Abstract** Urbanization is considered as the essentials of the modern world. The process of expanding the urban bodies in time and space is called as urbanization. It is a universal observation that all the urban expansions have engulfed almost all the water bodies in their territories and in the peripheral regions. Depletion of surface water bodies has put lot of adverse impacts on the water supply, groundwater development, and local climate. In a country like India where agriculture is based on the unorganized sector, farmers depend on groundwater and nearby water bodies for irrigation. Pollution and improper wastewater management have degraded the quality of surface water as well. Water availability has traditionally been monitored and assessed for various planning using expensive and time-consuming field techniques. The aim of this paper is to develop a model that could monitor how urbanization has affected the surrounding water bodies in the city of Madurai without going onto the field. This analysis works by overlaying existing geo-referenced data into a machine learning algorithm. For this project, we are using surface maps, groundwater maps, urbanization maps, and agriculture maps. Based on our designed prediction algorithm, we can analyze the gathered data, thereby creating a tool which can quantify the water stress, surface and groundwater depletion, and projected water availability.

**Keywords** Geo-referenced data · Qgis · AWEI · CART · Linear regression

## 1 Introduction

Water is a crucial resource, and in an agricultural country like India where most of the crops are rainfed ensuring availability becomes critical with irregular rainfall, drought, and increasing population. India is the wettest country in the world having average annual rainfall of 300–650 mm. However, its water resources are

---

This work was supported in part by the Department of Computer Science and Engineering of Thiagarajar College of Engineering

---

V. Vignaraj Ananth · R. Sarunisha (✉) · S. Aditya · M. Haripriya · M. Bhuvaneshwari  
CSE, Thiagarajar College of Engineering, Madurai, India

unevenly distributed both temporally and spatially and increasingly over-exploited by the human population.

Madurai, being a flat fertile plain, has its primary irrigation resource to be river Vaigai. When it comes to monitoring and managing water resources, the accuracy of monitoring results should not be compromised by the lack of proper validation. Meeting all of these criteria requires adequate financial and human resources. The field-survey methods used earlier were both expensive and time-consuming. This has been overcome in recent years after the intervention of various geo-spatial methods. The machine learning approach is employed to identify the relationship between state variables of the system and to overcome the time-consuming manual process with processes that are highly efficient and have a better accuracy. Machine learning techniques have proven efficient in predicting the non-stationary behavior of water resource systems. The sample size has a significant effect on the accuracy of the prediction. The objectives of this paper are to develop a methodology of monitoring the water resources available through spatial data and to assess using the automated water extraction index (AWEI) and predict the availability of the resource monitored.

Section 1 describes the introduction to effort estimation. Section 2 provides the literature survey. Section 3 shows the implementation. Section 4 describes the experimental results, and Sect. 5 provides the conclusion and future.

## **2 Basic Concepts**

### **2.1 Study Area**

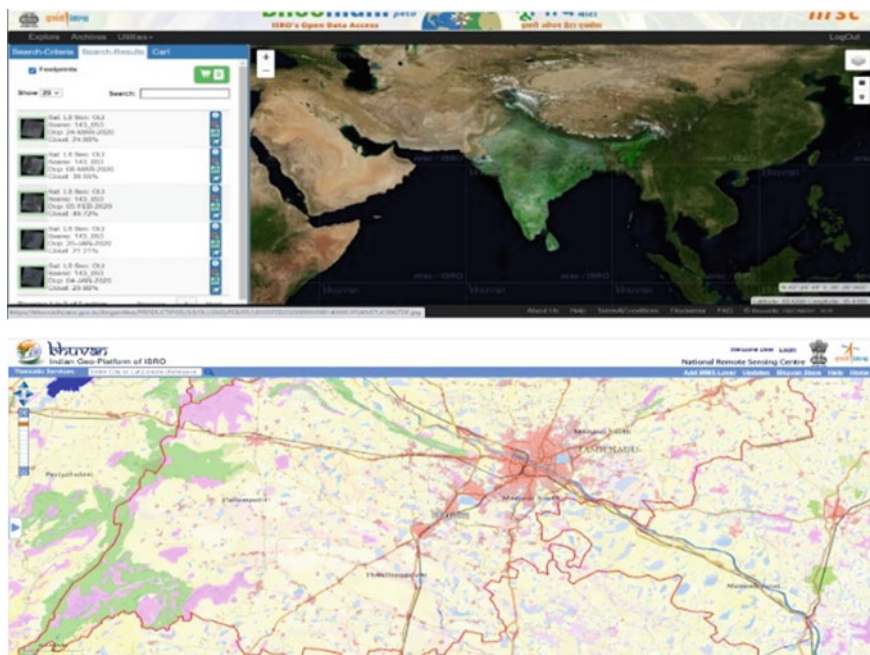
The area under study is the Madurai District in the state of Tamil Nadu.

### **2.2 Dataset**

The dataset includes the Landsat 8 OLI images of Madurai region along with the vegetation map and water fraction map downloaded from Resourcesat-1 are used to overlay and collect input data for the machine learning model.

### **2.3 Data Acquisition**

Researchers use various forms of spatial data to analyze the factors monitored. To monitor the water resources of any particular area, there are several features like temperature, vegetation, and rainfall that need to be monitored for accurate predictions. We have chosen the Landsat 8 satellite images that are available within an



**Fig. 1** Data download using Bhuvan

interval of 16 days and have huge data archives to analyze past trends. There are many sites to get these Landsat images, and we can derive it from BHUVAN the ISRO open data archive or also from USGS or through the Qgis software which collects data from USGS data catalogue. The data downloaded needs to be stored in separate folders along with metadata for further processing. The data is collected from 2015 to the current date to analyze the impact and potential of water bodies (Figs. 1, 2, and 3).

## 2.4 Preprocessing

Mostly the spatial images obtained are multilayered and have atmospheric disturbances. The data to be fed into any machine learning model need to be cleaned up and standardized. By doing so we can reduce the complexity of the learning model, and also, accuracy of the model can be increased. The preprocessing algorithm should be selected in such a way that the image does not lose any information. The first step is to clean the image. The entire processing of the image is done using the Qgis software. The atmospheric correction is done using the semiautomatic classification plugin (Fig. 4).



Fig. 2 Selecting data criteria and fixing the topo tile to be downloaded

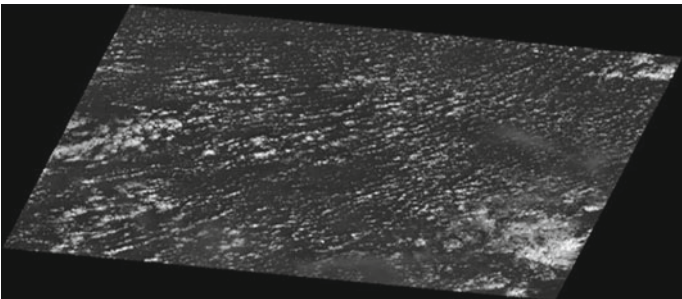


Fig. 3 Sample data images acquired from USGS of Landsat 8

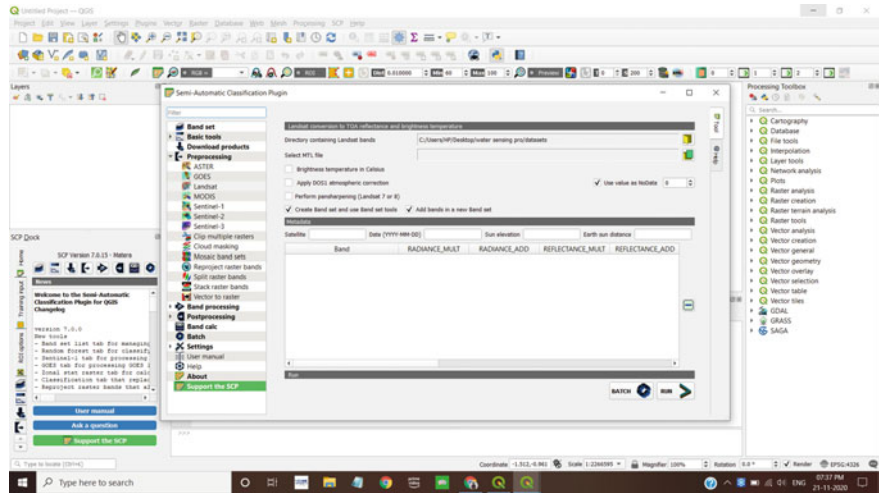


Fig. 4 Atmospheric correction using Qgis

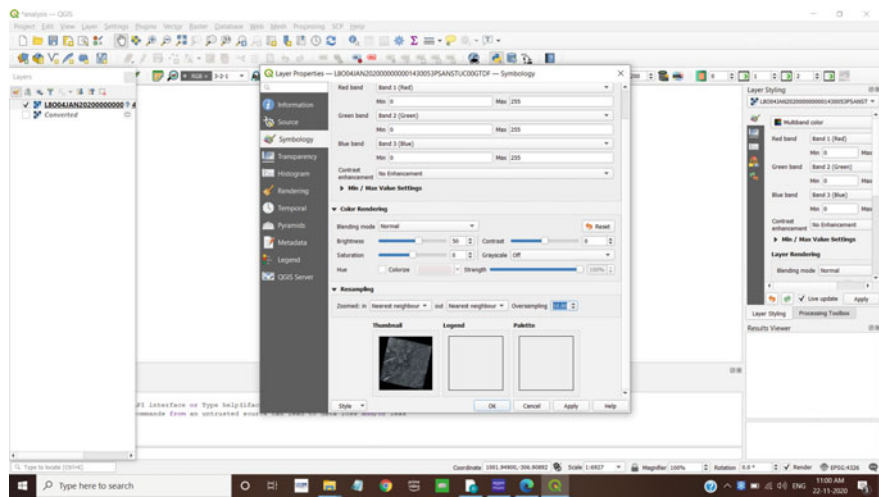


Fig. 5 Processing the data to get volume and other factors

2.5 Analyzing the Water Potential

The processed image is now again analyzed, and several other plugins are run; the water sources are marked polygons, and the volume calculator is used to find the water available. To avoid errors over wet areas or pits, analysis over the time series. The same processing is done to all batches of images, and the value found is tabulated. The output statistics help us to find the degradation of water level, and the usage trends can be plotted over time. The supporting factors as mentioned earlier are also monitored by running the specific plugins; hence, the temperature and vegetation cover are tabulated simultaneously (Fig. 5).

3 Implementation

3.1 Visualization of Dataset

See (Figs. 6, 7, 8, 9, and 10)

3.2 Algorithm Implementation

- (i) The classification and regression tree were used to develop the model that can classify the variables and henceforth identify the water resources highly

	A	B	C	D	E	F	G	H
1	Year	Month	Temp	Rainfall	Evapotranspiration	Obs. NDVI	Volume	
2	2015	jan	24.8	0	0	0.55	305	
3	2015	feb	26.89	0	0	0.54	308	
4	2015	mar	29.07	0	0.3	0.51	300	
5	2015	apr	31.87	0	1.5	0.49	302	
6	2015	may	32.48	0	2.3	0.54	301	
7	2015	jun	31.88	100.1	6.14	0.5	312	
8	2015	jul	31.52	199.4	5.69	0.5	320	
9	2015	aug	31.55	229.1	3.48	0.54	330	
10	2015	sep	31.04	136.7	1.17	0.55	328	
11	2015	oct	28.1	665.3	2.14	0.55	335	
12	2015	nov	25.67	-36.9	1.45	0.57	334	
13	2015	dec	29.9	-9.1	0.65	0.63	333	
14	2016	jan	26.94	-9.1	0.1	0.68	295	
15	2016	feb	29.72	-7.1	0	0.67	280	

Fig. 6 Dataset acquired through processing

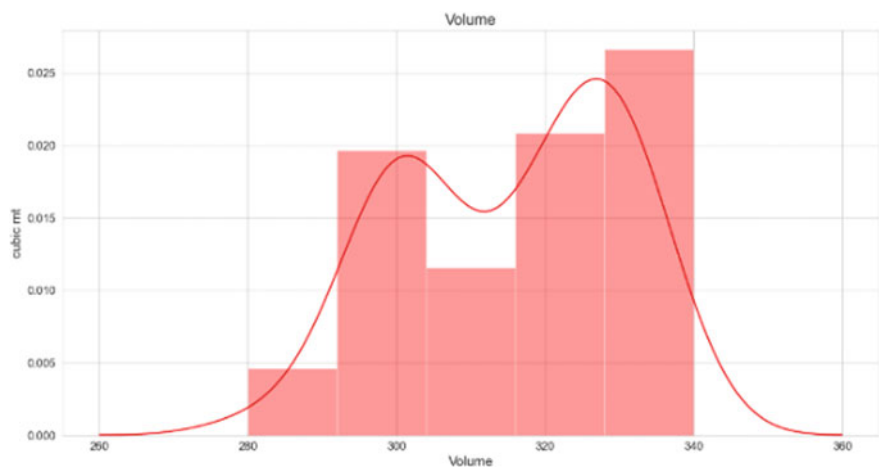


Fig. 7 Distribution plot of the dependent variable volume

- affected by the factors considered from all the water resources studied. For developing the model, the total sample was split into 80–20%. Development of the model uses 80% of the total sample data, and 20% is used as validation data.
- (ii) The initial classification starts with splitting the water bodies that are highly rainfed then using the interdependent factors of temperature, soil evapotranspiration, and vegetation. The Gini values are calculated for the instances considering minimum and maximum values of the variable, and this is repeated after every split in the tree.

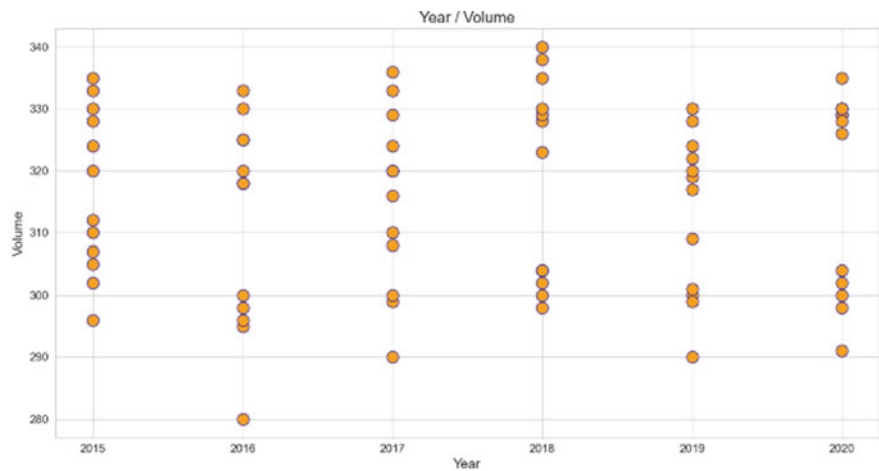


Fig. 8 Scatter plot of volume over the years

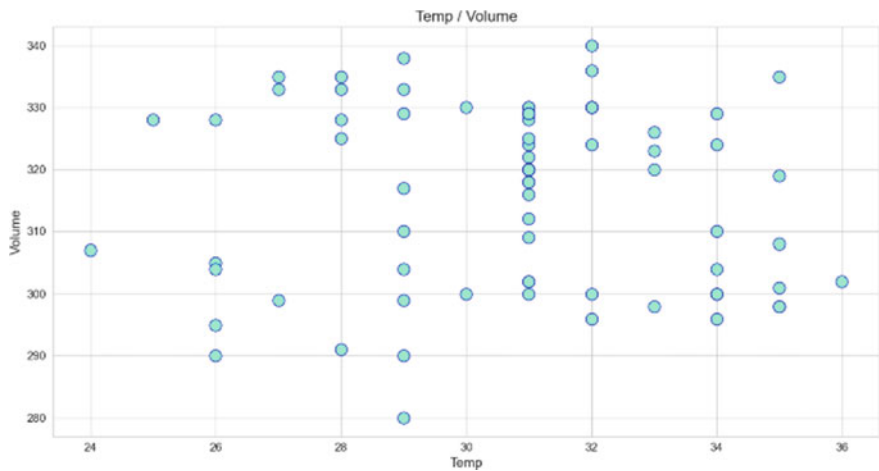
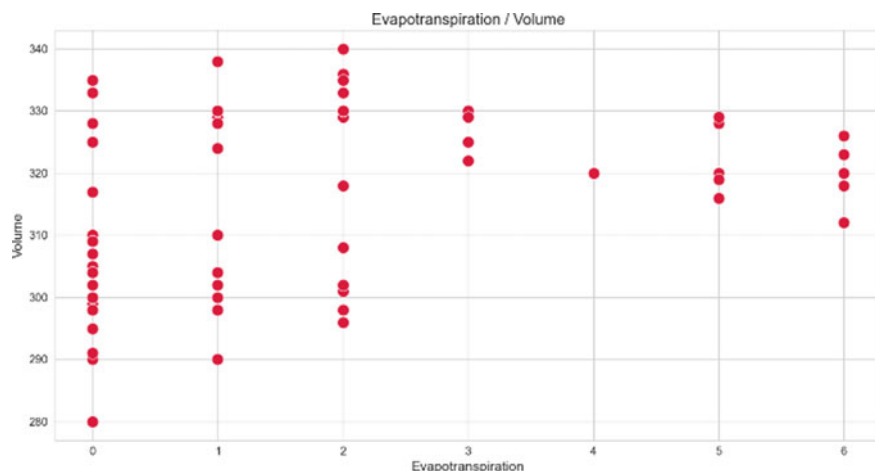


Fig. 9 Scatter plot of volume with temperature in the y-axis

- (iii) The CART method is able to determine the complex interactions among variables in the final tree. The final split determines the number of water bodies that we can rely on in the future derived based on the extremes of the given input (Fig. 11).





**Fig. 10** Scatter plot of volume with the varying evapotranspiration rate in the y-axis

```
# Create Decision Tree classifier object
clf = DecisionTreeClassifier()

# Train Decision Tree Classifier
clf = clf.fit(X_train,y_train)

#Predict the response for test dataset
y_pred = clf.predict(X_test)
```

**Fig. 11** Decision tree implementation

## 4 Experiments and Results

### 4.1 Model Validation

The mentioned models have implemented the variance, and R-squared values are discussed to find the fit and efficiency of the dataset to our model.

**Variance** is a measure of how far observed values differ from the average of predicted values. And the explained score is above average and proves the data enough to validate the model (Fig. 12).

Based on the CART and other regression models implemented using the derived dataset, we can analyze and infer the relation between variables and hence predict the range of values to sustain the model in the future (Fig. 13).

The obtained R-squared value is as follows: The least performing model is the Bayesian regression and the best fit for our dataset is found to be the ordinary least



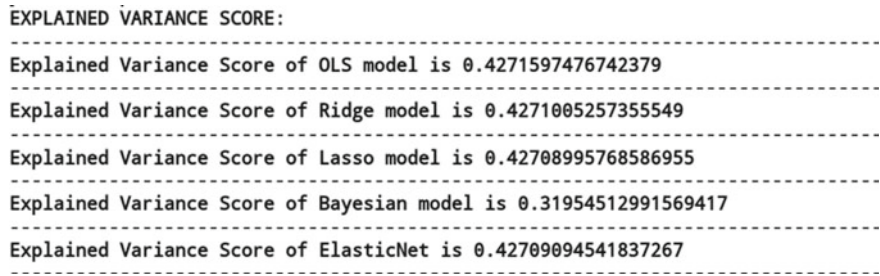


Fig. 12 Output variance score

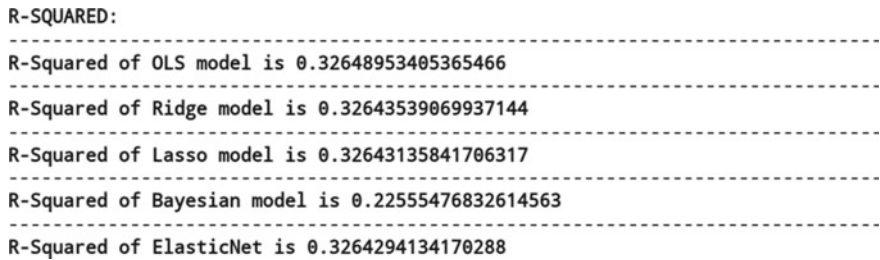


Fig. 13 R-squared values for the implemented models

squares. As we know the fit of the dataset does not define the efficiency of the model but ensures the adequacy of deterministic coefficient.

The processing of the Landsat 8 OLI images to identify the volume of the surface water resources takes into account all available water bodies, and the quality of water is not considered and the water resources reported of pollution and not utilized for any domestic or agricultural purpose is exempted and not used in the training model.

4.2 Spatial Distribution of the Water Bodies

The processing of the Landsat 8 OLI images to identify the volume of the surface water resources takes into account all available water bodies, and the quality of water is not considered and the water resources reported of pollution and not utilized for any domestic or agricultural purpose is exempted and not used in the training model.

5 Conclusion

This paper shows the importance of considering the climatological factors while monitoring and assessing the impact of urbanization on surface water. The use of the

Qgis, semiautomatic classification plugin for removing the atmospheric hindrances in the Landsat 8 images. With very limited access to the GIS or climatological data and the recent published reports being till 2015/16. It is crucial to study the water bodies for various affecting factors to predict the future availability. This methodology can be used on several other interdependent factors to predict the patterns. The insights out of this model can aid to sustain the resources and improve utilization of the available resources, thus providing a sustainable water management system benefitting the ecosystem.

Models and techniques for monitoring the surface water dynamics and the quality are extensively required. Although various algorithms have been applied for monitoring the surface water dynamics CART predictive (supervised) algorithm seems to provide a better insight on the dependency of the variables and hence providing a threshold value during the split of the variables for the water body to sustain. And this algorithm being nonparametric adds to the advantages while developing the model with variables that are interdependent.

# Applications of Open Source Intelligence in Crisis Analysis—A COVID-19 Case Study



A. K. Buvanasri, R. Meenakshi, and S. Karthika

**Abstract** The ongoing Coronavirus pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), also known as COVID-19, was first identified in the city of Wuhan in late December 2019. It was declared a pandemic during March 2020, as it caused widespread infections and deaths. COVID-19 has caused millions of deaths across the world, making it one of the deadliest pandemics that has ever occurred. The great pandemic affected billions of lives economically and emotionally. Traditionally, in-person surveys were organised in order to perform crisis analysis to identify and create statistics of people affected during such pandemic outbreaks or disasters. Being an extensive process by itself, the amount of time these surveys consume and the requirement of a large workforce to collect such statistics do not favour the government or the governing authorities in coming up with solutions and timely services that would upbring the affected set of people. Apart from taking into account the physical damage caused, there is no mechanism that gives importance to the mental state of the masses in such situations. This study aims at performing crisis analysis based on the emotions/reactions exhibited by the people on the Internet. The collection of the entire dataset is done with the help of open-source intelligence tools. A monthly sentiment analysis is performed to compare the emotions, followed by performing sentiment analysis on the dataset collected from March 2020 to March 2021 to analyse the emotions shown by people during COVID-19 and sentiment prediction with projected accuracies on the collected dataset that is done using six machine learning classification algorithms.

**Keywords** Open-source intelligence · Sentiment analysis · COVID-19

---

A. K. Buvanasri (✉) · R. Meenakshi · S. Karthika

Department of Information Technology, SSN College of Engineering, Chennai, Tamil Nadu, India  
e-mail: [buvanasriak17017@it.ssn.edu.in](mailto:buvanasriak17017@it.ssn.edu.in)

R. Meenakshi

e-mail: [meenakshir17048@it.ssn.edu.in](mailto:meenakshir17048@it.ssn.edu.in)

S. Karthika

e-mail: [skarthika@ssn.edu.in](mailto:skarthika@ssn.edu.in)

## 1 Introduction

The ongoing Coronavirus pandemic caused by severe acute respiratory syndrome coronavirus 2(SARS-CoV-2), also known as COVID-19, was first identified in the city of Wuhan in late December 2019. It was declared a pandemic during March 2020, as it caused widespread infections and deaths. COVID-19 has caused millions of deaths across the world, making it one of the deadliest pandemics that has ever occurred [1]. Globally, there have been 127,289,043 confirmed cases, 2,785,365 deaths, and 72,200,063 recoveries, as per World Health Organisation WHO statistics of 29 March 2021 [2]. Apart from direct impact such as increase in number of cases and death, the pandemic has become a great source of anxiety and stress, causing mental health issues which are indirectly affecting people's relationships with the society. The COVID-19 pandemic has been termed as the most consequential global crisis since the World Wars. It has had a prominent effect on human beings all over the world. The economic impact of COVID-19 is very alarming. The big shift in the world economic market and share market has affected various countries. The outburst of the virus has not only affected the global economy, but has also created a mayhem in people's minds. Most health professionals who tend to overwork in the times of pandemic suffer high levels of psychophysical stress and stress traumatic disorder. The widespread impact of COVID-19 infection has created a strong need for discovering efficient analytics methods for understanding the flow of information and the development of mass sentiment in pandemic scenarios. While there are numerous initiatives analyzing care and recovery, healthcare, preventative, economic and network data, there has been relatively little importance given to the analysis of social media communications and people's personal levels. Twitter, Facebook, etc., are great resources to encapsulate emotions and thoughts of humans. During these difficult times, people have taken to the Internet to discuss their opinions, fears and insights on the pandemic.

Global events such as pandemics and disasters have adverse effects on the collective mentality of the society. Hence, it becomes vital for governing authorities to derive conclusions from research and analysis works that help understand the needs and requirements of the general public to come up with a precise and sustainable scheme that can benefit everyone. While in-person surveys focus on the physical health status of the affected masses, analysing the situational reactions exhibited by people on the Internet allows us to understand the situational mentalities of the masses. This study aims at providing a structured analysis of the sentiments exhibited by people on the Internet in reference to the global pandemic, COVID-19 along with predicting the nature of sentiments while consuming an exponentially small amount of time when compared to performing door-to-door surveys.

The objectives of the proposed system include building a dataset from all over the Internet using open-source intelligence tools, performing crisis analysis and visualising the types of effects/impacts the ongoing pandemic has had on the society, and using machine learning classification algorithms to predict the sentiments.

## 2 Existing Works

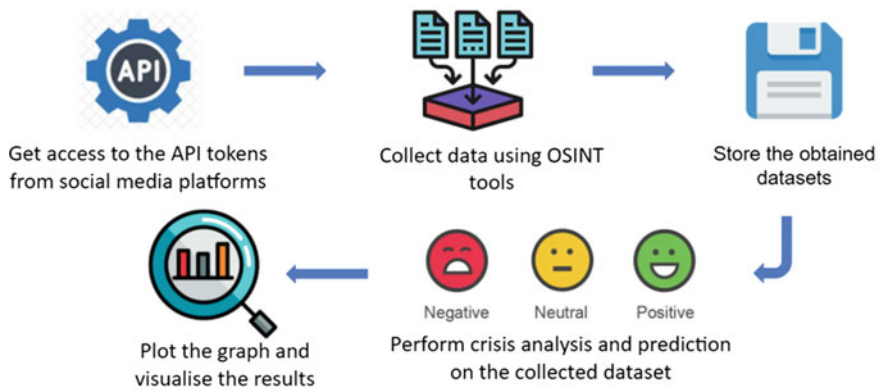
The concept of open-source intelligence (OSINT) is introduced as an important part of understanding the twenty-first century solutions to human problems in Glassman and Kang [3]. OSINT in many ways is the result of changing personal-knowledge relationships. The authors in Pastor-Galindo et al. [4] combine the future and the future of OSINT by analysing its positive and negative points, explaining ways to use this type of intelligence, and describing future indicators of the emergence of this process while those in the paper Backfried et al. [5] explore the role of open-source intelligence (OSINT) in disaster management. They discuss how open resources provide a cheap, fast and effective way to assess the situation of those affected in a disaster or in a time of crisis. The authors of Daia et al. [6] fixed their primary goal as to detect clusters via gensim libraries in news OSINT data that contains threat and health-related information.

The researchers in de Las Heras-Pedrosa et al. [7] analyse the way in which social media platforms have affected risk communication in the unsure contexts and the way in which it has affected people's emotions and sentiments in Spanish society during COVID-19 pandemic. A global sentiment analysis of Coronavirus-related tweets was done by the authors in Mansoor et al. [8] and changes in human emotions in different countries over time have been analysed. In the research work [9], the authors identified the public sentiment associated with the epidemic using specific Coronavirus and R-mathematical software, as well as its sentiment analysis packages whose insights are exhibited in the proceedings of fear-sentiment over time as COVID-19 is approaching the highest levels in the United States, using descriptive text analytics supported by the required textual data recognition.

## 3 Proposed System Design

In this section, the overall design of the system is discussed. The system has the following major modules namely data collection, sentiment analysis and machine learning prediction models used for classification of the processed data and visualisation of the analysed data in Fig. 1.

There are numerous open-source intelligence available on the Internet but finding the right tool plays a major role in collection of data that is relevant to the analysis. After an in-depth analysis of the existing OSINT tools, four tools showed potential output that can be used for this analysis. The data is collected from all over the internet using Open Source Intelligence tools, namely Social Bearing, Social Searcher, Twint and NetMiner. Social Searcher is a free social search engine that allows users to search across all major social media platforms. It measures and tracks mentions, presenting the results in an easy to use dashboard, which can be exported as a csv file. Social Bearing helps to analyze any Twitter account, hashtags, keywords, websites or users in a particular geographical location. It is advanced search toolbar allows



**Fig. 1** System architecture

users to filter results according to dates, language, exclude certain keywords or only see tweets from one user to the other and the report can be exported. NetMiner is a software for exploratory analysis and visualisation of enormous amounts of data. NetMiner’s Data Collector can collect data from various social media platforms such as Facebook, Instagram, Twitter, etc., with a couple of clicks. Twint is a complicated Twitter scraping tool written in Python that permits for scraping Tweets from Twitter profiles without using Twitter’s API. Twint uses Twitter’s search driver to scrape Tweets from any particular user, scrape tweets concerning certain topics, hashtags and trends. The various OSINT tools used for data collection are discussed below in Table 1.

Sentiment analysis (or opinion mining) is a natural language processing technique used to analyse whether a textual instance is positive, negative or neutral. Sentiment analysis models focus on polarity (positive, negative, neutral) but also on feelings and emotions such as anger, happiness, sadness, urgency, etc., and even intentions such as, if a person is interested or not interested in a particular scenario. Depending on what is needed, categories can be defined and tailored to meet the sentiment analysis needs. Sentiment analysis is extremely important in times of pandemic, such as COVID-19 as it helps the government to take the needful steps. It is performed

**Table 1** List of open-source intelligence tools

Tool name	Access	Link
Social bearing	Twitter	<a href="https://socialbearing.com/">https://socialbearing.com/</a>
Social searcher	All over the internet	<a href="https://www.social-sea-rcher.com/">https://www.social-sea-rcher.com/</a>
Twint	Twitter	<a href="https://github.com/twintproject/twint/">https://github.com/twintproject/twint/</a>
NetMiner	Twitter, Facebook, Instagram, YouTube	<a href="http://www.netminer.com/">http://www.netminer.com/</a>

using TextBlob, which is built on the shoulders of NLTK Python Library. TextBlob offers a lot of features like sentiment analysis, pos-tagging, noun phrase extraction, Tokenization, Words inflection and lemmatization. Tokenization refers to dividing text or a sentence into a sequence of tokens, which roughly correspond to “words”. Noun phrase extraction is particularly important when you want to analyze the “who” in a sentence. Various machine learning prediction models are used to train and test for the classification of the processed data, namely Naive Bayes, Multinomial Naive Bayes, Baseline, Input Mapped Classifier, Decision Stump and Random Forest. The visualisation of the processed dataset after sentiment analysis was done with the help of two open-source intelligence tools—Social Bearing and Social Searcher. The data was visualised by domain, by hour, by country, by number of posts per 5 s.

## 4 Data Preprocessing

Social Bearing is a Free Twitter analytics tool that searches for tweets, timelines and twitter maps. It helps to filter and sort tweets or people by engagement, influence, location and sentiment. The results can be exported as a csv file [10]. Social Searcher is a search engine that allows the user to monitor all public social mentions in social networks and the web. It helps to quickly measure and track what people are expressing about any topic, or event in one easy to use dashboard, which would be a huge time saver. The report can be exported as a csv file [11]. Twint is an advanced Twitter scraping and OSINT tool written in Python that doesn't use Twitter's API, allowing you to scrape a user's followers, following, Tweets and more while evading most API limitations [12]. NetMiner is a software for exploratory analysis and visualisation of enormous amounts of data. NetMiner's Data Collector can collect data from various social media platforms such as Facebook, Instagram, Twitter, etc., with a couple of clicks [13].

A total of 1 Lakh data instances were collected from March 2020 to March 2021 using the open-source intelligence tools. Moreover, two separate datasets were collected between the periods of 15, January 2021 to 14, February 2021 and 15, February 2021 to 14, March 2021 using open-source intelligence tools to perform monthly sentiment comparison. The data collected from various tools were in different formats. The common fields from all the datasets of different open-source tools were collected and combined to build the final dataset that was used for sentiment analysis. The data collected from various tools had factors such as timestamp, username, mentions links, which were irrelevant to the current analysis. The cleaning of data included removing those factors and it was done with the help of simple python scripts.

Data collected from all over the internet may be noisy, inconsistent and incomplete which can produce misleading results in data mining. Data preprocessing is a proven method of resolving such issues. It filters out useless data, and it also transforms raw data into an understandable format. Various steps involved in it are

**Table 2** Example from the dataset and their normalised versions

Raw	misses Swimming Class, <a href="http://plurk.com/p/12nt0b">http://plurk.com/p/12nt0b</a>
Normalized	misses swimming class URL
Raw	©98PXYRochester HEYYYYYYYYYY!! its Fer from Chile again
Normalized	USER_MENTION hey its fer from chile again
Raw	Sometimes, You gotta hate #Windows updates
Normalized	sometimes you gotta hate windows updates
Raw	@Santiago_Steph hii come talk to me i got candy :)
Normalized	USER_MENTION hii come talk to me i got candy EMO_POS
Raw	@bolly47 oh no :( r.i.p. your bella
Normalized	USER_MENTION oh no EMO_NEG r.i.p your bella

data cleaning, feature reduction, integration and transformation. Table 2 shows an example of differences that exist between raw and cleaned instances of text.

Sentiment analysis was performed on the overall dataset with the help of various python libraries, from March 2020 to March 2021, to find the emotions shown by the people during the times of COVID-19. A Monthly sentiment analysis was performed on two other datasets, 15, January 2021 to 14, February 2021 and 15, February 2021 to 14, March 2021 to compare and contrast the emotions. The Monthly analysis was done because, in the months of January 2021, news about COVID-19s mutation was spreading among people. The comparison helps to determine if the COVID-19s mutation affected the people.

## 5 Results and Discussion

The evaluation metrics such as positive, negative and neutral can be found. The scores 0, 1 and 2 indicate the sentiments neutral, negative and positive, respectively. The major difference between other text classifiers and the classifiers used in this study is the usage of the word “positive” and “negative” in interchangeable scenarios as “tested positive for COVID-19” is a negative instance while “tested negative for COVID-19” is a positive instance here. Hence, necessary changes were made to the collected dataset manually so as to yield better accuracies from the machine learning training models. Figure 2 shows the sample of the classification of textual instances.

### 5.1 Monthly Comparison of COVID-19 Emotions

Datasets belong to the periods 15, January 2021 to 14, February 2021 and 15, February 2021 to 14, March 2021. An analysis is performed on both datasets followed



text	score	sentiment
chaotic situation puerto ricans indignant tourists breaking covid mandates via assembly m j rivers death covid 2 workers	1	negative
piotr radwan1 piotrwjtowicz6 gov pl cancer lecture unfortunately covid regardless stress movement state health know healthy athletic person unfortunately died quit c	1	negative
driving upturn activity sales business covid 19 plus star startups making happen ai startups	0	neutral
far anti vaxxer tony excuse wanting inject something body little research done regards long term ramifications taking said vaccine ve covid ll let	2	positive
ridiculous writes incredibly helpful books fact crib sheet number one resource mom don t agree take kids covid fine stop attacking providing opinion different	2	positive
thank cripps full house covid vaccines dashed surgery vaccine shift amp efficient well done amp team	2	positive
repairs cases structure covid 19	0	neutral

Fig. 2 Sample of the classification of sentiments as positive, neutral and negative

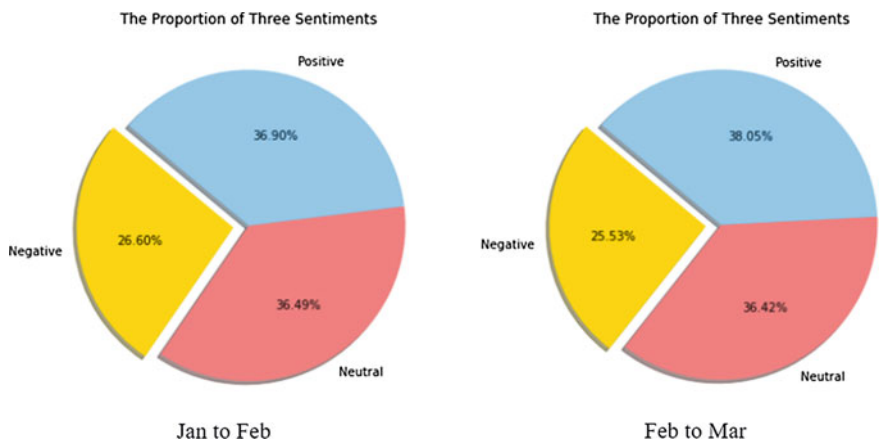


Fig. 3 Comparison of sentiment distribution

by comparison of the same. From looking at the pie charts in Fig. 3, there is a slight increase in percentage of positive sentiments and decrease in percentage of negative sentiments in comparing the datasets of Jan-Feb and Feb-Mar which shows the effect of commencement in distribution of vaccines against COVID-19.

In Fig. 4, some words like “Trump” and “Americans” have been mentioned. It indicates the happening of US Presidential elections in the month of January amidst COVID-19 pandemic.

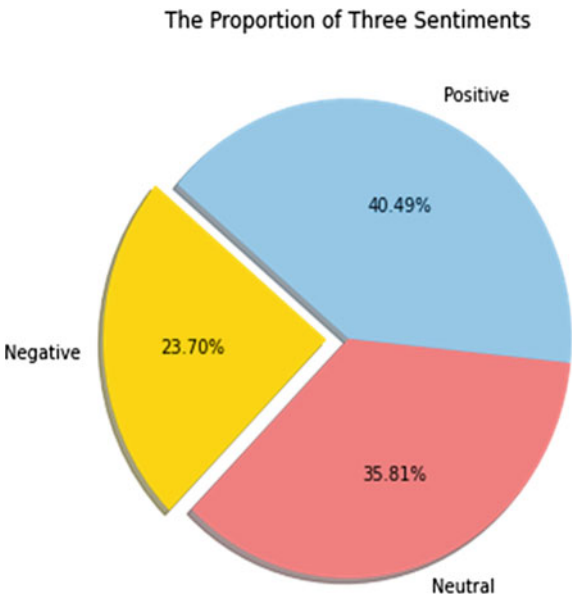
5.2 Overall Analysis of COVID-19 Emotions

Figure 5 depicts the percentage split of positive, negative and neutral sentiments



Fig. 4 Comparison of word clouds

Fig. 5 Pie chart of sentiment distribution



during the period of March 2020 to March 2021. There is a fair share of positive and neutral sentiments while there is comparatively less amount of negative sentiments. Figure 6 depicts the frequency of the most common words such as “covid”, “cases”, “deaths”, “pandemic”, “vaccine,” etc., found in the collected dataset in the period of one year between March 2020 and March 2021.

5.3 Accuracy Comparison of Classification Models

Machine learning models are used against the dataset that has over 90,000 cleaned textual instances which has been collected over the pandemic period between March 2020 and March 2021 of which 80% of it is used for training the models and the

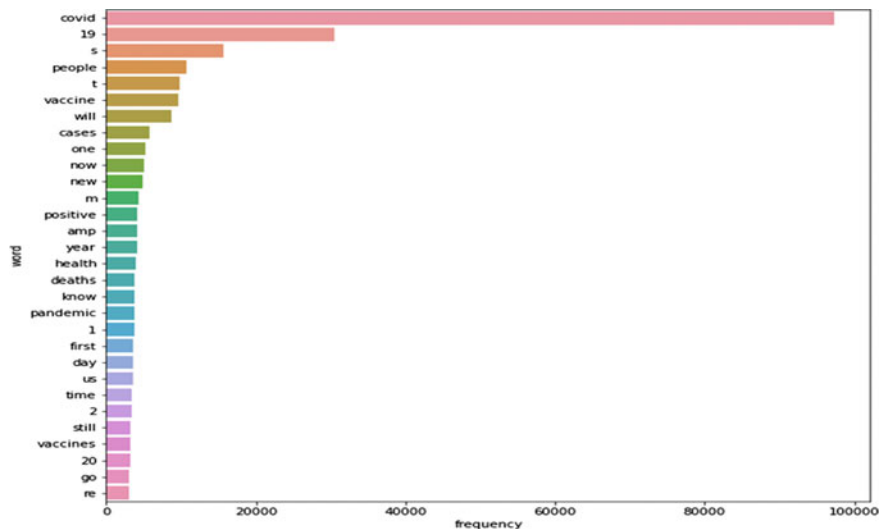


Fig. 6 Snapshot of frequency of words

remaining is used for testing the accuracies of the predicted sentiments. The training dataset has text instances labelled as positive, negative and neutral which were added manually. The various models used in this study are Decision Stump, Random Forest, Input Mapped Classifier, Baseline, Naive Bayes and Multinomial Naive Bayes.

The accuracy of the Decision Stump algorithm is 40.55% as, for continuous features, a threshold feature value is selected, and the stump contains only two leaves which are for the values below and above the threshold. Hence, the accuracy is reduced. Random Forest consists of a large number of individual decision trees that operate as an ensemble. The predictions and the errors made by the individual trees need to have low correlations with each other. The accuracy is 38.60% as the correlation between each model is high. The Input Mapped classifier yields an accuracy of only 40.55% as it works based on Decision Stump (here), and it can be seen that even with additional attributes in test data, the same accuracy as to Decision Stump is obtained as only the necessary attributes viz., cleaned text instances, its sentiment and its sentiment score are considered. The accuracy of the baseline algorithm is 32.73% as it finds the k-many closest labels for the given data instance. The accuracy of Naive Bayes algorithm is 60.81% because, if a categorical variable has a category (in the test data set), which was not observed in the training data set, then the model will assign a 0 (zero) probability and will be unable to make a prediction. Multinomial Naive Bayes uses one time occurrence of the words to contribute to the results instead of using a full term frequency. This is the reason as to why the Multinomial Naive Bayes algorithm when trained gives better results compared to other classification algorithms used. It gives a maximum accuracy of 77.00%.

Table 3 shows the proportion of data used for training and testing, along with

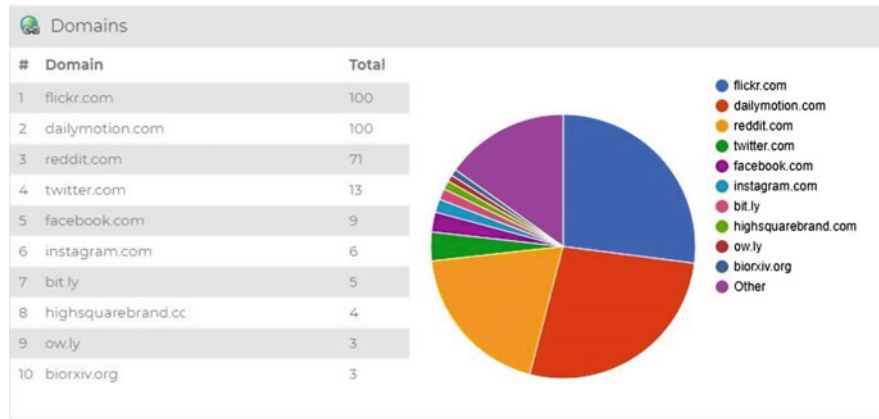
**Table 3** Accuracy of the classification algorithms

Algorithm	Dataset proportion		Accuracy (%)
	Trainin (%)	Testing (%)	
Input mapped classifier	80	20	39.7
Decision stump	80	20	40.5
Random forest	80	20	30.6
Baseline	90	10	32.7
Naive bayes	90	10	60.8
Multinomial Naive bayes	90	10	77.0

the outcome of each algorithm where the lowest accuracy of 32.7% is marked by baseline algorithm as it finds the k-many closest labels for the given data instance while the maximum accuracy of 77.00% is obtained using Multinomial Naive Bayes algorithm as it looks for the occurrence of positive or negative words instead of their frequencies. Hence, it can be concluded that the Multinomial Naive Bayes algorithm is suitable for textual sentiment prediction.

5.4 Data Visualisation

The Data visualisation of the collected dataset was done with the help of OSINT Tools—Social Bearing and Social Searcher. The data visualisation is done by domain, hour, country and posts per 5 s. The data is collected from various domains. The top 10 domains that contributed to the data collection are shown in Fig. 7. It can be seen



**Fig. 7** Data visualisation by domain

that flickr, Dailymotion, reddit and twitter are major contributors of the dataset. The same has been visualised as a pie chart. Figure 8 shows the number of posts that are being uploaded on the internet for every 5 s. It can be noted that almost 10 posts are being uploaded in 5 s.

Figure 9 shows the time at which users are most active to post on various social media platforms in the times of COVID-19. The dataset is collected from all over the world with a maximum upload of data of about 39% which has been happening between 4 and 6 AM according to the Indian Standard Time. Figure 10 shows the countries from which the dataset is collected. The countries are coloured from light green to dark green, with dark green being the country with the most number of users

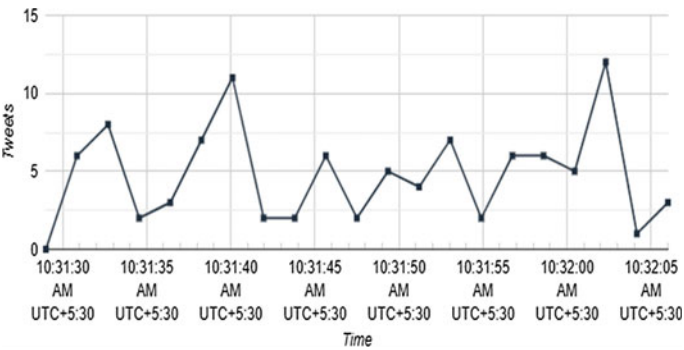


Fig. 8 Data visualisation by posts per 5 s

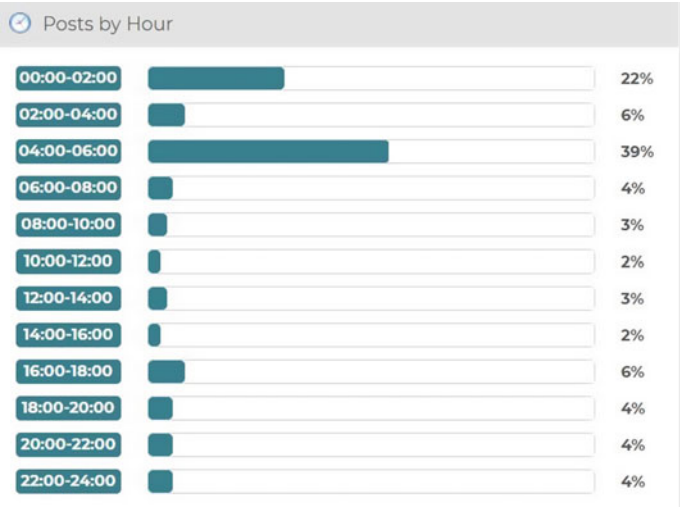
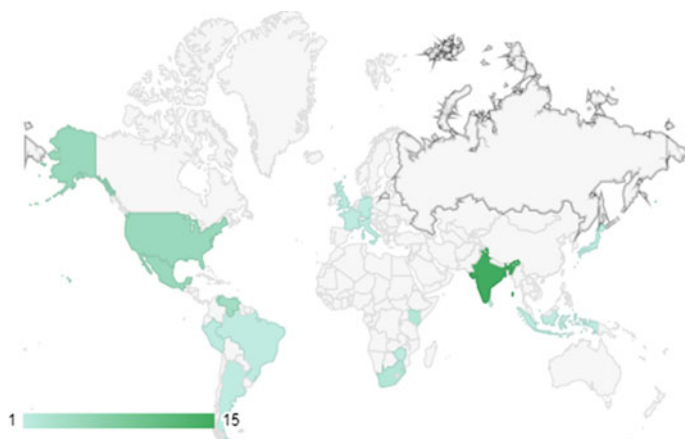


Fig. 9 Data visualisation by hour



**Fig. 10** Data visualisation by country

and light green being the countries with least number of users. It can be seen that India has the most number of users.

## 6 Conclusion

The world is in the midst of a pandemic, COVID-19. Covid-19, a novel virus, continues to infect people all over the world despite the discovery and distribution of vaccines. As of 18 February 2021, at least seven different vaccines across three platforms have been rolled out in countries [14]. This study focuses on the usage of social media analysis with the help of OSINT tools for data collection rather than social network analysis. It also utilises the potential of social media analysis in understanding as well as classifying scenarios in case of a pandemic or a natural calamity. Though the conventional methodologies of crisis analysis tend to endure such difficult situations, one can take advantage of the fact that the generation of information on social media platforms increases day-by-day, to assess the needs of individuals during unexpected scenarios such as COVID-19. It has been found that although scientific research works on COVID-19 are growing tremendously, there is space for open source data collection and extraction in multiple directions. This study collects a dataset of approximately 1 lakh textual instances from various online platforms, performs sentiment analysis and analyses the performance of various classifying models. After extensive analysis, it was observed that Multinomial Naive Bayes gave the highest accuracy of 77%. This study makes a substantial contribution in analysing the emotions of the people and how COVID-19 has affected them mentally. Hence, ruling authorities can defer to efficient analytical methods that take advantage of technological leaps, to truly understand the effect and nature of such physically unobservable scenarios.

Future work of this study would involve using more OSINT tools for the collection of data and the sentiment analysis of people towards other topics such as mental health, government response to the pandemic, healthcare facilities, etc. It is also recommended to apply feature selection and parameter tuning while building machine learning models, and experiment using deep learning models.

## References

1. Shuja J, Alanazi E, Alasmay W, Alashaikh A (2020) Covid-19 open source data sets: a comprehensive survey. *Appl Intell* 1–30:1
2. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>. Last accessed 29 Mar 2021
3. Glassman M, Kang MJ (2012) Intelligence in the internet age: The emergence and evolution of open source intelligence (OSINT). *Comput Hum Behav* 28(2):673–682
4. Pastor-Galindo J, Nespoli P, Mármol FG, Pérez GM (2020) The not yet exploited goldmine of OSINT: opportunities, open challenges and future trends. *IEEE Access* 8:10282–10304
5. Backfried G, Schmidt C, Pfeiffer M, Quirchmayr G, Glanzer M, Rainer K (2012) Open source intelligence in disaster management. In: 2012 European intelligence and security informatics conference. Odense, Denmark, pp 254–258. <https://doi.org/10.1109/EISIC.2012.42>
6. Daia A, Stancu S, Vladoi A, Ionescu-Tîrgoviște C (2020) Clustering analysis on news from health OSINT data regarding CORONAVIRUS-COVID 19. *New Trends Sustain Bus Consumption* 669
7. de Las Heras-Pedrosa C, Sánchez-Núñez P, Peláez JI (2020) Sentiment analysis and emotion understanding during the covid-19 pandemic in Spain and its impact on digital ecosystems. *Int J Environ Res Public Health* 17(15):5542
8. Mansoor M, Gurumurthy K, Prasad VR (2020) Global sentiment analysis of COVID-19 tweets over time. *arXiv preprint* [arXiv:2010.14234](https://arxiv.org/abs/2010.14234)
9. Samuel J, Ali GG, Rahman M, Esawi EK, Samuel Y (2020) Covid-19 public sentiment insights and machine learning for tweets classification. *Information* 11(6):314
10. Social bearing. <https://socialbearing.com/>. Last accessed 25 Mar 2021
11. Social searcher. <https://www.social-searcher.com/>. Last accessed 26 Mar 2021
12. NetMiner. <http://www.netminer.com/main/main-read.do>. Last accessed 10 Mar 2021
13. twint. <https://github.com/twintproject/twint>. Last accessed 26 Mar 2021
14. Covid-19 Vaccines. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines>. Last accessed 27 Mar 2021

# Language Agnostic Generation of Syntactically Incorrect Programs



Ramyani Ghosh, Rasya Ramesh, V. S. Meghana, and N. S. Kumar

**Abstract** In recent times, migration to the online world has been at the forefront, highlighting the importance of online education. This paper illustrates the development of an automated online tool for novice programmers. Every programming language is governed by its syntax and semantics. Mastering the syntax of a language is the first step toward mastering programming in that language. We have proposed a multipronged tool called “CodeLearn” which assists in generating syntactically incorrect programs, tests a learner’s ability to correct them and also produces a dataset of grammatically wrong programs from a given correct program. We focus on the generation of syntactically incorrect programs by injecting meaningful errors into grammatically correct programs. Furthermore, the generation of errors is made language agnostic, implying that we can generate erroneous programs for any programming language, given its grammar. We expect that this tool could help students, teachers, and research scholars. Teachers can generate personalized question papers with debugging questions for students. Researchers can generate a dataset of incorrect programs for a given piece of code and use this corpus to reverse engineer and find the right program, i.e., facilitate automatic syntax error correction.

**Keywords** Program synthesis · Formal grammar · Education · Smart learning · Compiler design · Programming languages

## 1 Introduction

A fundamental concept in computer science is the learning and understanding of programming languages. As a novice, we take our first steps in the world of computer science by looking at a high-level view of various problems. We try to understand the working of different algorithms to find solutions to these problems. After all, the final goal of learning this subject is to solve problems. However, one often finds oneself in

---

R. Ghosh (✉) · R. Ramesh · V. S. Meghana · N. S. Kumar  
PES University, Bangalore, Karnataka, India

N. S. Kumar  
e-mail: [nskumar@pes.edu](mailto:nskumar@pes.edu)



a situation where even with an understanding of the big picture, a lot of time is wasted in trying to adjust to the syntax of the language used to solve the problem. In this section, we describe the importance and means of learning syntax. In the following sections, we elaborate on our novel approach to tackle the problem of syntax learning by developing a language agnostic algorithm that generates debugging exercises for novice programmers. Following this, we illustrate a UI in the form of a tool we developed called “CodeLearn,” which provides various features that utilize our algorithm.

### ***1.1 Importance of Learning Syntax***

The first step in the learning of languages is to understand the syntax of the language in question. Without a sound understanding of syntax, hindrance free writing of code is not possible, and debugging is bothersome.

### ***1.2 Facilitating Learning of Syntax***

It is common for a teacher to simulate syntax errors and give debugging exercises to help students deepen their understanding of a language. This derives from the teaching idea of “learning from mistakes”. This is when a student learns by recognizing the difference between a wrong answer and a correct one, and subsequently correcting one’s mistakes. Inspired by this learning process, we introduce error correction exercises for learning programming languages. The outline of the exercises is as follows:

1. A teacher presents a source code with some syntax errors in it.
2. Students must find all the faults in the presented code.
3. Students must fix the wrong code.

As the program is not being written from scratch by the student, it is a great beginning step in the process of understanding the syntax of a language with little prior knowledge.

## **2 Related Work**

Katsuyuki et al. worked on developing a debugging exercise extraction system [1] using learning history. They developed an editing history visualization system. It consists of learning logs of incomplete programs and accumulates the edit history of the programs. The tool compares all the previous versions with the latest versions,

thus extracting the number of mistakes. A machine learning approach was also implemented by Rahul Gupta et al. in their paper where they generate a novel programming language correction framework [2] amenable to reinforcement learning. The tool goes through each line and corrects the error. Rewards are given at each action of the tool. When it receives the maximum reward, it is assumed that the tool has successfully corrected the code. These papers have explored syntax errors from a programmer's point of view whereas our work focuses on the novel applications of introducing syntax errors in programs.

### 3 Parsers in Compiler Design

In order to structurally interpret a program, we need to convert the program into a data structure. A parser breaks down each statement of the program into units which can then be organized into a parse tree. This is done using compiler tools such as lex and yacc. In our implementation, we use Python Lex and Yacc (or PLY.), Beazley [3] to parse the input code segments, produce a parse tree, then modify this tree to generate error ridden programs.

#### 3.1 *Elements of Python Lex and Yacc*

**Tokens.** The tokens are a list of legal symbols or patterns permitted by a particular language. These tokens include a variety of common syntax categories such as keywords, operators, identifiers, etc. Each of these types are declared in distinct lists. A token may be defined in structure such as in the case of keywords, or it could be in the form of a regular expression- setting the rules for all allowed strings.

**Action Functions.** Grammar productions are denoted by action functions in PLY [4]. These functions contain the grammar rule indicating which tokens can be replaced by which expressions. These functions enforce that an input can be parsed only if the language rules are satisfied and there are no syntax errors. These functions also contain the vital operations that pertain to building the parse tree at every stage.

**Other functions.** Some other functions defined in the PLY program that are crucial for gleaning all the required information from the parsing process include:

1. Function that keeps track of new lines and line numbers.
2. Function to indicate a break in the parsing flow, i.e., a syntax error due to the violation of a grammar rule.
3. Function to ignore whitespaces and tabs that do not have an effect on the syntax.
4. Function to check for reserved words and prohibit them from being considered as identifiers.

## 4 Implementation

### 4.1 *Generating PLY Program for Language Agnosticism*

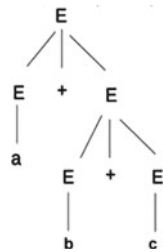
Making a program language agnostic means to make it compatible with multiple possible languages which can be taken as input. To parse inputs given in various languages, our parsing functions in PLY have to be adaptive to a language. To do this, we provide a structured grammar (syntax rules and productions) for any language to our program generation code, which automatically converts these rules to PLY functions. These PLY functions can parse any given code in the language of the input grammar, given that the input code follows the syntax of the language precisely. Thus, we are synthesizing a PLY program that will be different for every programming language. This process enables language agnosticism. This generation of a PLY program is carried out by taking the grammar for a language in a text file as input and collecting all relevant data from it. Once we have generated the PLY program with specific token lists and action functions, we can parse all input code segments in that language and build a parse tree, as described later in this paper.

### 4.2 *Error Injection*

Generating errors while randomizing the process and generalizing it for any given code segment and language is a non-trivial task. We go about this by following these steps:

**Parsing Input Code.** The tool runs the generated PLY program on the input code segment for the given language. If the input code segment has no syntax errors, it runs the whole segment through the parser, and each token is assigned a node in the parse tree. At the parsing process's termination, we successfully assemble a parse tree that contains every token present in the input code. The tool can subsequently flatten this tree to give us a continuous string that is basically the input program.

**Fig. 1** Possible parse tree for “ $a + b + c$ ” generated from grammar  $G$



**Modifying Parse Tree.** Generating errors in the program can be done quite simply by making three kinds of modifications to the parse tree for the program: addition, removal, and substitution. This can be illustrated with the help of a simple example. Consider the following grammar  $G$ :

$$\begin{aligned} E &\rightarrow E + E \\ E &\rightarrow a \\ E &\rightarrow b \\ E &\rightarrow c \end{aligned}$$

where  $E$  is the starting symbol, and  $\{a, b, c, +\}$  is the set of finite terminals, that is, the tokens of the language. The given grammar, although ambiguous, can be used to produce a parse tree for the following input:

$$a + b + c$$

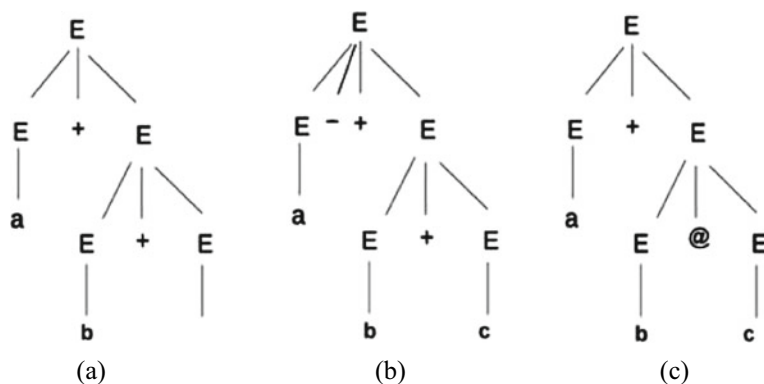
A possible parse tree that can be constructed for this string from the given grammar is illustrated in Fig. 1.

The yield of the parse tree is nothing but the program segment itself (*In our case:*  $a + b + c$ ). Thus, by manipulating the yield, errors can be generated in the program segment. We introduce errors by making the following three types of modifications to the parse tree produced:

*Removal of a node.* The first way to introduce an error into the program would be to remove a node from the leaves of the tree. For example, consider the tree in Fig. 2a. Here, the token “ $c$ ” has been removed, thus making the resulting program segment:

$$a + b +$$

*Addition of a node.* The second tactic would be to insert a node into the leaves of the tree. We have considered certain metrics to perform this insertion, which are specified in Sect. 4.3. For example, consider the tree in Fig. 2b. Here, the token “ $-$ ” has been inserted, thus making the resulting program segment as follows:



**Fig. 2** Parse trees after different methods of error generation

$$a - +b + c$$

*Replacement of a node.* The third method is to replace one of the existing nodes with another similar one in the same location (Fig. 2c). For example, replacing a “+” with “@” in the parse tree, gives the resulting erroneous program segment:

$$a + b@c$$

### 4.3 Error Generation Algorithm

Our error generation algorithm goes through the parse tree and chooses a node at random, and then arbitrarily performs one of the following operations:

1. Replacement with a new node in accordance with the heuristics table (see Table 1).
2. Addition of a new node right after the chosen node by either duplicating it or adding a new and relevant value from the list of chosen values. For example:

$$\text{for}(i = 0; i < n; i++)$$

3. Removal of the chosen node.

**Table 1** Error generation rules

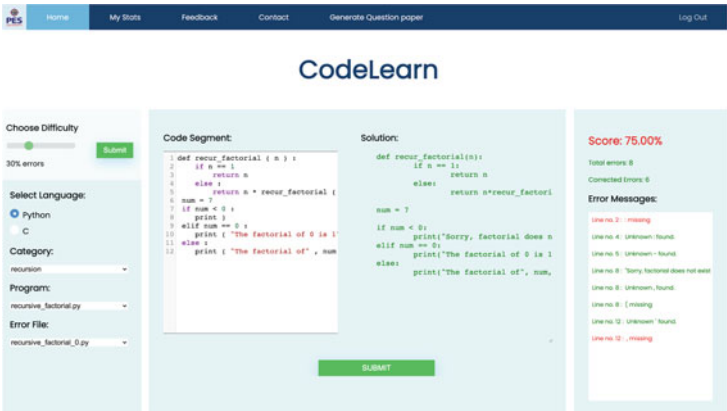
S No.	Error generation algorithm	
	Section being swapped with	Section swapped into
1	Brackets: ( { }	Brackets: ( { }
2	Selection keywords: if switch else	Selection keywords: if switch else
3	Loop keywords: for while	Loop keywords: for while
4	Equality operators: < = > =	Symbols: ; : .,
5	Arithmetic operators: * - +	Symbols: ; : .,
6	Symbols: ; : .,	Symbols: ; : .,
7	Reserved keywords: void, int	Misspellings: viod, itn

5 Application

We use this error generation algorithm to develop a learning tool for novice programmers. Our major application is in the form of our website “CodeLearn” as explained subsequently.

5.1 User Interface

We have designed a tool for both teachers and students. Our user interface is in the form of a website we developed called the “CodeLearn Tool.” The UI contains the following features:



**Fig. 3** User interface of CodeLearn tool

**Auto-evaluated Debugging exercises for Students.** We present the student with a tool that displays error programs in an editor (see Fig. 3) to correct the syntax errors and receive a score. The components of the page include:

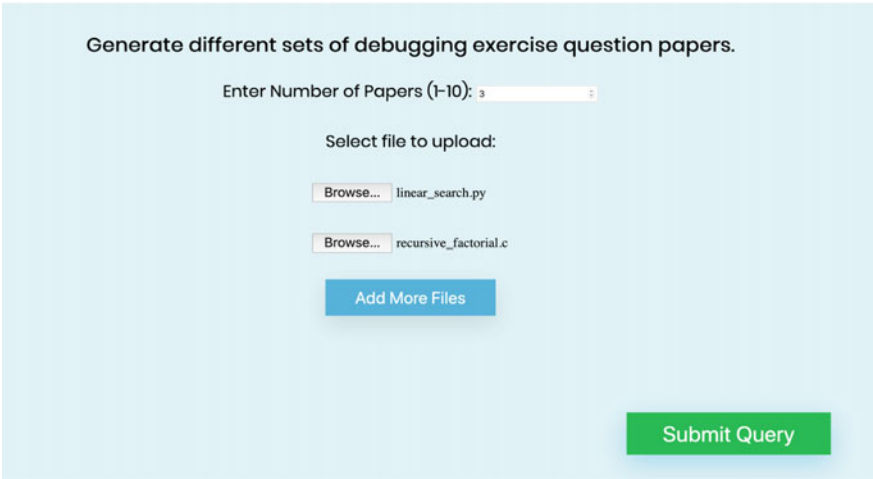
1. Dropdown options to choose a programming language.
2. Choice of programs to edit under various categories such as data structures, arrays, recursion, etc.
3. A percentage slider for the user to choose the difficulty level of the code by sliding the error percentage bar from 0 to 100 percent. Higher the number of errors, the more difficult it becomes to detect all of them.
4. A user-friendly text editor that displays the erroneous code where the student can edit and correct the errors by typing.
5. A submission button for the changed code which then displays a score based on the number of errors successfully found and corrected by the student.
6. A line wise report of corrected and missed errors and the relevant error messages (Fig. 3).
7. A “View Solution” option, in case the student wants to skip the exercise and view the solution.

## 5.2 Dataset Generation

Our program makes multiple corresponding erroneous programs for each correct input program and stores them in a repository. Thus, our dataset contains correct programs mapped to multiple corresponding error programs. Other data stored for each entry includes program category (such as recursion, searching, sorting, etc.) and the percentage of errors injected into the code (with respect to the total number of nodes in the equivalent parse tree for the input program). This dataset can be used as a repository of syntactically correct programs mapped to syntactically incorrect programs, feeding a potential machine learning program that can potentially reverse our process to automatically correct syntax errors in a program.

## 5.3 Question Paper Generation

Another feature of our CodeLearn Tool is that a teacher can generate question papers with debugging exercises (see Fig. 4). The teacher is asked to upload files containing correct programs. Our tool then parses these files (dependent on the programming language) and injects errors into them. Thus, we generate multiple sets of question papers containing different erroneous versions of the input code. Each set will have different errors in the same code. The teacher can choose the number of sets of question papers required and then download the generated papers with the debugging exercises.



**Fig. 4** Question paper generation for teachers

## 6 Evaluation and Results

We tested our error generation program using our user-friendly interface. We passed some of the error programs generated through the actual compilers for the given languages to test whether the programs generated were actually syntactically incorrect. To evaluate the accuracy of our error generation, we use a metric called Syntax Error Validity (1):

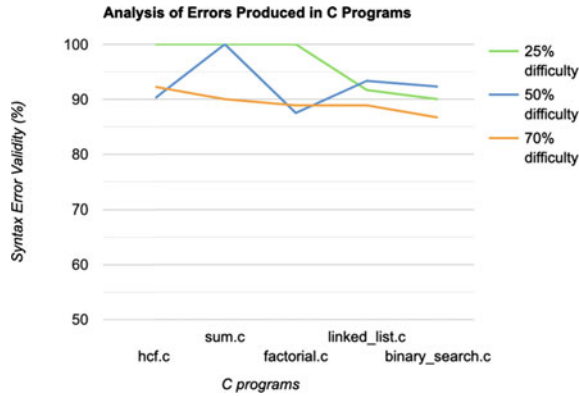
$$\frac{\text{Number of accurate syntax errors}}{\text{Total number of changes injected}} * 100 \tag{1}$$

Thus, we used this metric to plot the performance of our error generation algorithm. We see that with greater difficulty level (more errors introduced) and for larger programs (such as linkedlist.c), the number of inaccurate errors introduced is more. These errors include semantic errors or changes that don't result in a syntax error (for example: a double semicolon). The results of our findings are plotted in the graph (see Fig. 5).

The debugging tool has been tested by multiple people (including fellow students) for the purpose of feedback. It seems to perform the desired functionality fairly successfully. Novice programmers have said that this is a helpful exercise for them to perform online, without the need for a teacher.



**Fig. 5** Evaluation of error generation accuracy



## References

1. Umezawa K, Nakazawa M, Goto M, Hirasawa S (2019) Development of debugging exercise extraction system using learning history. In: IEEE tenth international conference on technology for education (T4E), pp 244–245
2. Rahul G, Kanade A, Shevade S (2018) Deep reinforcement learning for programming language correction. In: arXiv preprint [arXiv:1801.10467](https://arxiv.org/abs/1801.10467)
3. Beazley D (2001) Ply (python lex-yacc). <http://www.dabeaz.com/ply>
4. Beazley D (2007) Writing parsers and compilers with PLY. In: PyCon'07, Texas

# Distributed FBG Temperature Sensor for Coal Mine Fire Detection



Sagupha Parween and Aruna Tripathy

**Abstract** In underground coal mines, fire is one of the most horrible hazards. Underground coal mine fire mostly occurs due to smoldering combustion. The smoldering combustion occurs by self-heating due to an increase in temperature because of exothermic internal reactions leading to explosion. The temperature gradient is one of the most important factors for detecting underground fire. The forthcoming mine fire always leads to gradual rise in temperature. Therefore, detection of temperature change in underground coal mines is an important concern. The detection can be done by using conventional temperature sensors like thermocouple, resistance temperature detector (RTD)s, and thermistors. An optical fiber Bragg grating (FBG) sensor is better than these electrical sensors as it is immune to electromagnetic interferences. We have proposed and demonstrated an FBG temperature sensor based system, where multiple FBG sensors sense different temperatures, and the sensed data is transmitted through a single channel by using wavelength division multiplexing (WDM). Multiplexing reduces the number of components and installation cost within a sensor network. At the receiver side, a WDM demultiplexer is used to retrieve the originally transmitted signal. The performance of WDM-FBG sensor system has been compared in two different mediums namely; wired and wireless. For wired communication, optical fiber cable (OFC) has been used and for wireless communication, free space optics (FSO) channel has been used.

**Keywords** Optical temperature sensor · FBG · Bragg wavelength · WLS · WDM · OFC · FSO · Coal mine fire detection

## 1 Introduction

Underground coal mines have difficult working conditions, and there are many hindrances encountered during the process of coal mining. The coal mine fire is

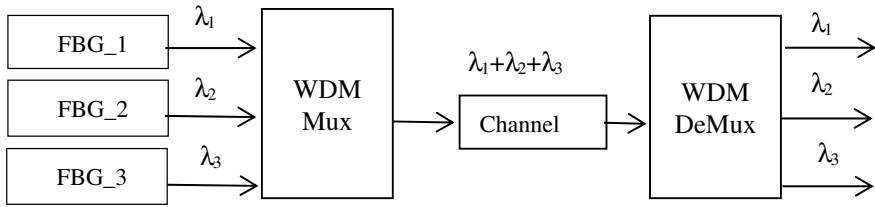
---

S. Parween (✉) · A. Tripathy  
College of Engineering and Technology, Bhubaneswar, Odisha, India

A. Tripathy  
e-mail: [atripathy@cet.edu.in](mailto:atripathy@cet.edu.in)

one of the terrible hazards which occur due to self-heating when coal oxidizes naturally over time and results in an exothermic reaction that generates heat [1]. The heat produced determines the energy transfer between the coal and its surroundings due to the temperature changes. The event occurs when coal absorbs oxygen in the air. The coal oxidizes and ends in a flaming fire caused by the accumulated heat [1, 2]. This fire leads to toxic gas emission, increment of temperature of surroundings, combustion, loss of precious coal, followed by increment of production cost which are some of the major consequences of mine fire. It is very much important to take safety measures before the initiation of this condition. The fire detection forthcoming parameters are temperature, gas, smoke, smell, humidity, air velocity, etc. In this paper, our main focus is on temperature detection. The coal mine fire leads to gradual increase in temperature; therefore, detection of temperature change in underground coal mines is an important concern. For this multiple fiber Bragg grating (FBG) sensors can be used for detecting deviation in temperature in underground coal mines to detect the coal mine fire before its combustion. Conventionally, electrical temperature sensors like thermistors (range:  $-50$  to  $250$  °C), resistance temperature detectors (RTD) (range:  $-196$  to  $482$  °C), and thermocouples (range:  $-250$  to  $2320$  °C) have been used [3] for various applications. Optical sensors have many advantages over electrical sensors such as high sensitivity, small size, no electromagnetic interference, and light weight [4]. These sensors can work with a small transmission power which is a boon for hazardous operating conditions inside a coalmine. FBG mechanism is one of the most popular choices for optical fiber sensors for a wide range of temperature detection. The fiber core has periodic structure having different refractive indices which are called as grating [5]. The optical spectrum given by a white light source (WLS) is passed through the grating. At each grating period, a slice of wavelength is reflected due to change in reflective index inside the core and remaining parts are transmitted [6]. The combination of all the sliced wavelengths to one large reflection and the wavelength at which this reflection occurs is called the Bragg wavelength [7]. When WLS is given as an input to the FBG sensor, any change in temperature with respect to the reference temperature will lead to reflect back the center Bragg wavelength and pass the rest of the wavelengths. For the detection of gradual change in temperature in coal mines, a sensor network is required where many sensor nodes are distributed in a required area to sense or detect any change in environmental conditions. The channel capacity of transmission for a sensor network should be increased to transmit multiple sensed data simultaneously with high data rate at a time from different allocated sensors, and this can be done by using wavelength division multiplexing (WDM) [8, 9]. The basic block diagram of WDM system is shown in Fig. 1.

The transmitters  $T_{X1}$ ,  $T_{X2}$ , and  $T_{X3}$  shown in Fig. 1 are individual FBG transmitted data. The output signal from these transmitters at corresponding wavelength is multiplexed by a WDM multiplexer. The wavelength multiplexer combines the entire output signal, and the combined output is transmitted over a channel to reach at the receiving end [9]. The channel can be either wired or wireless. At the receiver section, the combined transmitted signals are separated and de-multiplexed by a



**Fig. 1** Block diagram of WDM-FBG sensor system

wavelength division demultiplexer (WDM DeMux) and the respective receivers  $R_{X1}$ ,  $R_{X2}$ , and  $R_{X3}$  receives their awaited data signal.

We have demonstrated the use of FBG sensors in a 8-channel WDM transmission system. The working of an FSO is similar to OFC with the only difference being the optical signal is sent through free air in the absence of optical cable. The light beams travel in free space [10]. The FSO has two units, i.e., transmitter and receiver unit each having optical trans receiver which provide full duplex connection [10]. The transmitter unit transmits invisible infrared light beam from transmitter and focuses it to the receiver in line-of-sight (LOS) condition [11]. The analysis of WDM-FBG system is carried out over both wired and wireless transmission channels, i.e., optical fiber cable (OFC) and free space optic (FSO) channel respectively. The range of the communication channels is also found and compared by changing the transmitted power of the source for both the OFC and FSO medium.

## 2 FBG System Model

An FBG is created by designing periodic patterns of change in refractive index inside the fiber core. This change in refractive index is typically created by showing the fiber core to an intense interference pattern of UV energy [11]. This will produce a permanent increase in the refractive index of the fiber's core, creating a fixed index modulation according to the pattern. This fixed index modulation is called a grating [11]. This fixed index modulation defines the grating period. At each grating period, a small amount of light is reflected. All the reflected light signals combine to one large reflection at a particular wavelength. This wavelength is stated as the Bragg condition, and the wavelength at which this reflection occurs is called the Bragg wavelength [12]. Only those wavelengths are reflected that satisfy the Bragg condition, and this wavelength has maximum efficiency. The Bragg wavelength is given by [12],

$$\lambda_B = 2n_{\text{eff}}\Lambda \quad (1)$$

where  $n_{\text{eff}}$  is the effective refractive index,  $\Lambda$  is the grating period, and  $\lambda_B$  is the Bragg wavelength. From Eq. (1), we can see that Bragg wavelength will shift with respect to any change in effective refractive index or grating period. The effective

refractive index has a value as average of all the periodic refractive indices in the optical fiber [13]. As the Bragg wavelength depends upon effective refractive index  $n_{\text{eff}}$ , any change in effective refractive index will shift the Bragg wavelength. Change in temperature range depends upon the value of thermo-optic coefficient of FBG Sensor. The change in wavelength with respect to temperature is given by [14],

$$(\Delta\lambda_B)_{\text{temp}} = \lambda_B(1 + \xi)\Delta T \quad (2)$$

where  $\Delta\lambda_B$  is the wavelength shift,  $\lambda_B$  is the center Bragg wavelength,  $\xi$  is the thermo-optic coefficient, and  $\Delta T$  is the change in temperature.

### 3 Proposed Model

Distributed sensors can be used to detect the temperature change in different places in coal mines. The WDM-FBG system transmits multiple sensed data simultaneously with high data rate at a time from different sensor over a communication channel to the receiver. This paper illustrates transmission of the sensed data over both a wired and a wireless channel. For wired channel, OFC is used and for wireless, FSO channel has been used to transmit the sensed data to the receiver. The schematic block of 8-channel WDM-FBG system with OFC and FSO as communication channels is shown in Figs. 2 and 3, respectively.

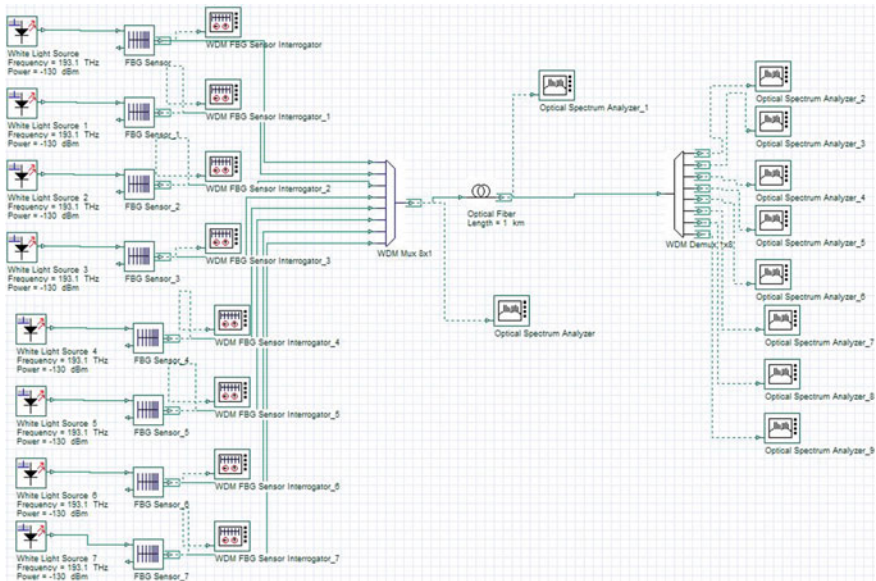
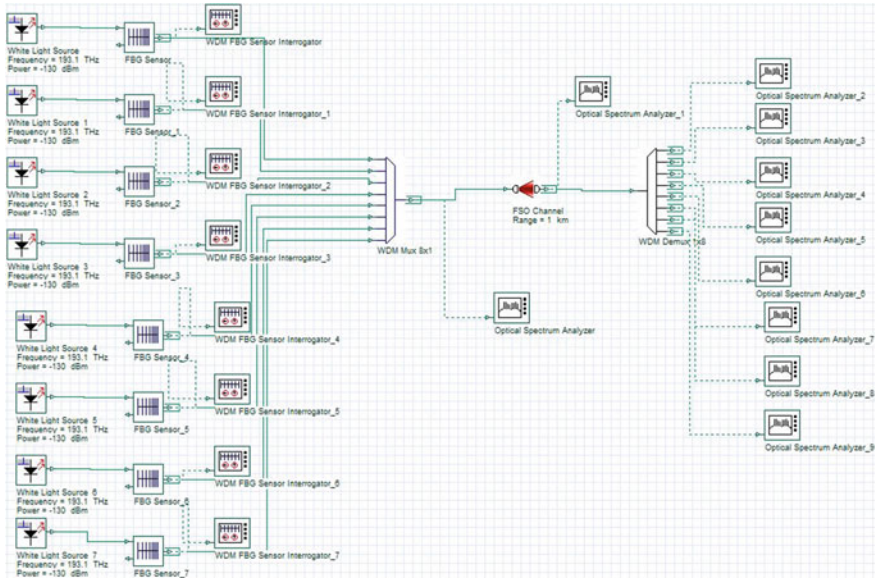


Fig. 2 Layout of the 8-channel WDM-FBG system over OFC channel

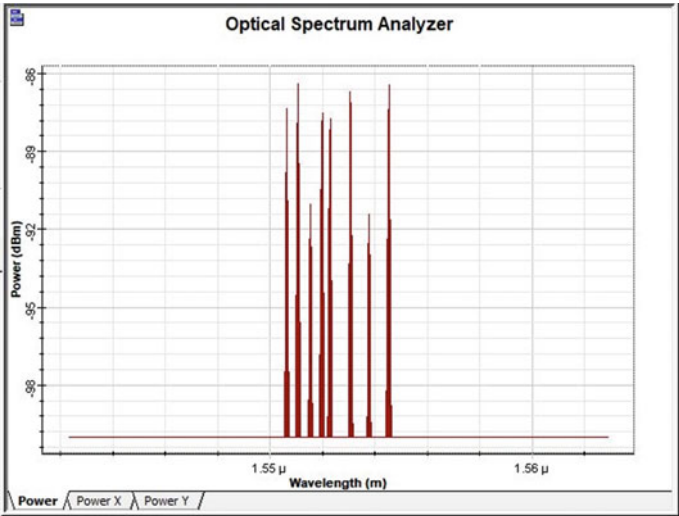


**Fig. 3** Layout of the 8-channel WDM-FBG system over FSO channel

In Figs. 2 and 3, a WLS is given as an input to each FBG sensor. In case of any deviation of the temperature in the surrounding from a reference temperature level, the center Bragg wavelength is shifted as given in Eq. (2). Those wavelengths that satisfy the Bragg condition are reflected, and rest is transmitted. The output spectrum can be seen through WDM-FBG sensor interrogator. WDM sensor interrogator is a tool for locating maxima or minima of the spectrum.

## 4 Results and Discussion

All the experimental block schematics are implemented in OptiSystem 16.0, and the comparative plot is simulated using MATLAB R2016b. In 8-Channel WDM-FBG system, eight number of FBG sensors that are used to sense 8 different temperatures, i.e., between 50 and 300 °C with respect to a reference temperature of 0 °C that has a center Bragg wavelength 1.54999 nm and  $\xi$  value 0.00001. The change in temperature will shift the center wavelength into different wavelengths, and these wavelengths are multiplexed using WDM which is shown in Fig. 4 and passed through an OFC link having length 1 km with channel attenuation of value 0.2 dB/Km. At the receiver side, WDM demultiplexer is used to retrieve the data from each transmitter successfully at the receiver side. Similarly, we have changed the communication channel to FSO channel and carried out the above test.



**Fig. 4** The  $8 \times 1$  WDM mux output passed through OFC channel

Due to any change in the temperature from the reference temperature  $0\text{ }^{\circ}\text{C}$ , the sensed Bragg wavelength in 8-channel WDM-FBG system that has been found through simulation is given in Table 1, for both OFC and FSO channel. Also, the shift in theoretical and sensed Bragg wavelength with respect to reference center wavelength, i.e.,  $1.54999\text{ nm}$  has been compared and given in Table 1, and the comparative plot for the same has been shown in Fig. 5.

**Table 1** Comparative summary of theoretical and sensed Bragg wavelength shift in 8-channel WDM-FBG system and shift in Bragg wavelength sensed through OFC and FSO channel

T ( $^{\circ}\text{C}$ )	Bragg wavelength through OFC ( $\mu\text{m}$ )	Bragg wavelength through FSO ( $\mu\text{m}$ )	Theoretical ( $\Delta\lambda_B$ ) in (nm)	Simulated ( $\Delta\lambda_B$ ) over OFC in (nm)	Simulated ( $\Delta\lambda_B$ ) over FSO in (nm)
50	1.55063	1.5506	7.75	7.75315	7.753
70	1.55104	1.5510	10.8	10.8572	10.85
100	1.55154	1.5520	15.5	15.5154	15.52
130	1.55198	1.55222	20.15	20.1757	20.1788
150	1.55229	1.55228	23.25	23.2843	23.2842
200	1.55307	1.55304	31	31.0614	31.0608
250	1.55398	1.55408	38.75	38.8495	38.852
300	1.55453	1.55449	46.5	46.6359	46.6347

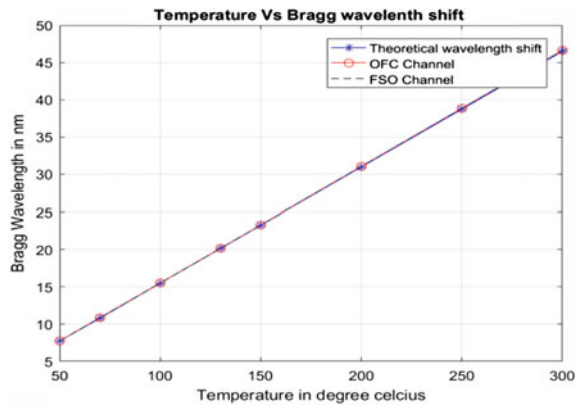


Fig. 5 Temperature verses Bragg wavelength shift

Table 2 Communication range of OFC and FSO channel for given transmitted power

Transmitted power (dBm)	Communication range in OFC (km)	Communication range in FSO
– 130	75	510 m
– 110	170	4 km
– 100	220	12 km
– 90	270	27 km
– 70	375	80 km
– 40	400	200 km
– 10	465	320 km

To receive the sensed data at the receiver side, the range of the communication channel should be known otherwise the data cannot be retrieved successfully. To increase the communication range, the transmitting power of each transmitter should be increased. The communication range of OFC and FSO channel is observed to be increased by increasing the transmitter power is given in Table 2. Beyond the given communication range in Table 2, the receiver is unable to retrieve the data.

5 Conclusion

In this paper, we have demonstrated an 8-channel WDM-FBG system. We have obtained the characteristics of FBG sensor by changing different values of the temperature for detecting the gradual increase in coal mine fire. The sensed data from FBG sensor has been sent through wired and wireless channel, i.e., through OFC and FSO communication channels. The Bragg wavelength shift through both OFC and FSO channel has been compared with the theoretically calculated Bragg wavelength



shift with respect to change in temperature, and it is found to be nearly same. This implies it is possible to send the sensed data from multiple FBG sensors through both wired and wireless channels. The communication range of OFC is up to 75 km and for FSO channel is up to 510 m at 0.2 dB/Km attenuation and transmission power of -130 dBm. We have increased the transmitter power up to -10 dBm and found the communication range for OFC is 465 km and for FSO is 320 km. But by increasing the communication range, the computational time is also increasing. To test the temperature ranges by FBGs in actual coalmines and to physically establish the channels to further test the validity of the obtained results as through simulation remains to be ascertained to establish the efficacy of the simulation models. Further, different optical sensors can be used to test the other parameters for fire detection in coal mines.

## References

1. Guansheng Q, Deming W et al (2016) Smoldering combustion of coal under forced air flow: experimental investigation. *J Fire Sci* 34:267–288
2. Onifade M, Genc B (2019) Spontaneous combustion liability of coal and coal-shale: a review of prediction methods. *J Coal Sci Technol* 6:151–168
3. Mohammed A, Babani S, Sanka AI, Abdullahi NA (2015) A comparative study between different types of temperature sensor. *Int J Ind Electron Electric Eng* 3:11–13
4. Jonathan O, Tobias M, Felix D, Konstantinos M, Johannes R, Klaus D, Roths J, Drechsler K (2018) Fiber optic measurement system for fresnel reflection sensing: calibration, uncertainty, and exemplary application in temperature-modulated isothermal polymer curing. *J Lightwave Technol* 36:939–945
5. Kadhim SA, Taha SAA, Resen DA (2015) Temperature sensor based on fiber bragg grating (FBG), implementation, evaluation and spectral characterization study. *Int J Innov Res Sci Eng Technol* 4:8038–8043
6. Ugale SP, Mishra V (2010) Fiber Bragg grating modeling, characterization and optimization with different, index profiles. *Int J Eng Sci Technol* 2:4463–4468
7. Hill KO, Meltz G (1997) Fiber Bragg grating technology fundamentals and overview. *J Lightwave Technol* 15:1263–1276
8. Ahmed BA, Aghzout O, Chakkour M et al (2019) Transmission performance analysis of WDM radio over fiber technology for next generation long-haul optical networks. *Res Article Int J Opt* 1:1–9
9. Rajan D (2016) Analysis of 8 channel WDM network With EDFA. *Int J Sci Eng Res* 7:1446–1449
10. Schroeder K, Ecke W, Apitz J, Lembke E, Lenschow G (2006) Fibre Bragg grating sensor system monitors operational load in a wind turbine rotor blade. *Meas Sci Technol* 17:1167–1172
11. Hong-Nan L, Dong-Sheng L, Gang-Bing S (2004) Recent applications of fiber optic sensors to health monitoring in civil engineering. *Eng Struct* 26:1647–1657
12. CE Campanella, A Cuccovillo, C Campanella, et al (2018) Fibre Bragg grating based strain sensors: review of technology and applications. *Art Multi Dig Publish Instit (MDPI) J* 18:1–27
13. Erdogan T (1997) Fiber Grating Spectra. *J Lightwave Technol* 15:1277–1294
14. Khare RP (2004) Fiber optics and optoelectronics. Oxford University Press

# Image Captioning Using Ensemble Model



Pranav Bathija, Harsh Chawla, Ashish Bhat, and Arti Deshpande

**Abstract** Generation of image captions is the task of producing a description of an image using natural language. Even with the advancement in technology, the mundane task of image captioning still proves to be strenuous for a machine. The paper proposes an ensemble model of LSTM and transformer encodings. The model works on an encoder–decoder architecture. The image is passed as input to the encoder which extracts features from the image which acts as input for the decoder to output a caption from the received information. The paper evaluates the results of the implementation of the ensemble model as well as the existing model. The paper also evaluates the results on different standards of images. The ensemble model provides better results than the standalone existing models. The results were evaluated on Bleu score. The proposed model achieved a bleu score of 0.6192.

**Keywords** Deep learning · Image captioning · Long short-term memory · Transformer · Natural language processing

## 1 Introduction

There is a quote “A picture is worth a thousand words”, a miniscule image contains a plethora of information. Humans are capable of processing such large amounts of information and comprehend it in an instant. Communication is an integral part of human existence, and humans use written or spoken languages as a medium of communication. Humans similarly use natural languages to describe an image. This also leads to multiple possible captions for the same image. This task proves to be much more complex for machines; however, if done successfully, it can prove beneficial for a variety of applications. The machine needs to have basic knowledge of natural languages, object detection and correlation to produce a caption for an image. Few of the previous models work on precise syntax and hard-coded features

---

P. Bathija (✉) · H. Chawla · A. Bhat · A. Deshpande  
Thadomal Shahani Engineering College, Mumbai, India

A. Deshpande  
e-mail: [arti.deshpande@thadomal.org](mailto:arti.deshpande@thadomal.org)

[1] to provide a caption for an image. As a result, the comprehensive nature of the caption generated is bounded by the previous approaches. Hence, the training data should be used as guidelines to build models that are free of sentence syntax or pre-defined features [1] to overcome this limitation.

Irrespective of the limitations, the task of generating image captions has displayed a range of possible use cases such as visual aid products for improving the life of visually impaired people. Augmented reality products such as Google lens can use it for semantic image search. It has also proved useful in bringing visual intelligence to chatbots and image tagging for social media, which with the recent surge in e-commerce makes it quite advantageous.

With the latest evolution in deep learning techniques, the feasibility to achieve a descriptive caption for an image has increased. The proposed model consists of mainly two components—encoder and decoder [2]. The encoder extracts important features from the image that acts as the input to the decoder which uses its knowledge gained during training to identify the objects individually and also find the relation shown between them in the image. The decoder is responsible to generate a caption from the word vector formed at the decoder using natural language processing.

This paper proposes a new model, which is based on an ensemble of two previous approaches, LSTM and transformer. The paper includes a study on the existing models that generate a caption for an unknown image as well as results of execution of the suggested ensemble model. The paper also aims to compare the proposed model and the existing models on images with clear and distinct objects, versus images with clusters and vague features. The F1 score of the proposed ensemble model shows a slight improvement from the results of the existing models solely based on either LSTM or transformer for the generation of the image caption.

The flow of the remaining paper is as follows. In Sect. 2, the paper discusses the recent research work conducted in this field. In Sect. 3, the dataset and the pre-processing techniques used are described. Section 4 discusses the design of the architecture of the proposed ensemble model. Section 5 evaluates the results achieved using the proposed model in comparison with the existing models. The paper concludes with direction for future work.

## 2 Related Work

Generation of image captions can be performed in mainly three different ways: template-based methods, retrieval methods and generative methods. In template-based methods [3–6], objects, actions, scenes, and attributes are prioritised to be detected. Im2Txt model [7] is one of the finest implementations of the retrieval-based method. However, the need to implement newer and better models arise out of the fact that retrieval-based methods only refer to their database images in order to rank content. This content which is ranked is then used to produce the captions. Hence, these methods are incapable of generating novel captions. The generative method uses either a pipeline or an end-to-end model [8] which makes collective

use of the language modelling and image recognition models. They are usually a combination of CNNs and RNNs.

In Shah et al. [1], a Show and Tell model is used for this implementation which combines Inception-v3 model and LSTM cells [2] where the former carries out object recognition, while the latter carries out language modelling [9, 10]. Bilingual evaluation understudy (BLEU) score is used to evaluate the generated sentences which work by comparing the generated sentences to human-generated sentences [11]. This implementation [1] gives an average BLEU score of 65.5. The paper [12] uses a neural framework where visual group geometry (VGG) network is used along with an LSTM network [13]. The VGG extracts the features from the images which then act as input to the LSTM network in the process of caption generation. Input to the LSTM network is an encoded fixed length vector of images and sentences. BLEU scores for the Flickr8k, Flickr30k and MSCOCO datasets are found to be 0.53, 0.61, and 0.67, respectively.

The author [14] does an analysis of the various image captioning types and the respective approaches taken in each type and provides a summary. The three types of image captioning systems mentioned are general image captioning, image captioning with emotion and cross-lingual and multilingual image captioning. General image captioning involves two processes: detecting the objects, their attributes, their relationships with other objects [15, 16] and modelling the appropriate language [17] which is semantically sound and best describes the image [18]. Whereas, captioning images in style by adding emotional content can be done by including the viewer's emotions and feelings towards the image [19–22] or by extracting emotions from the image itself [23]. Cross-lingual captioning can be done in several ways including direct translation into the desired language or training the dataset in the desired language. Apart from these three methods, [14] also discusses various datasets like UIUC PASCAL [24], Flickr 30 k [25] and Microsoft COCO Captions [26] alongside various evaluation metrics including BLEU [27], METEOR [28], ROGUE [29], CIDEr [30] and SPICE [31].

### 3 Dataset and Pre-processing

For the task of image captioning the data required is an image to be captioned as the input and a set of captions as the output to train the model. Several datasets exist which fulfil these requirements.

The most extensively used datasets are the Flickr30k and Flickr8k. Each dataset contains a lot of images with each accompanied by five captions which are used to train the model. The captions for these images are obtained via crowdsourcing. Figure 1 shows an image accompanied with its five captions from the dataset. The captions for the picture shown in the figure are as follows:

- A man wearing a red jacket is climbing a mountain.
- A person in red jacket standing on a snowy mountain.



**Fig. 1** Example image from Flickr30k

- A person standing on a snowy mountain.
- A man with a red backpack climbs a tall mountain.
- A man on a mountain.

Before training can begin, some pre-processing steps need to be performed. The captions need to be pre-processed so that they can be passed to the model. This is done by making all the captions lower case, removing all punctuations, special characters and converting numbers to their alphabetical representation. This is followed by tokenization of the captions. Since the model predicts the caption iteratively, the captions need to be converted in a format as shown in Table 1.

**Table 1** Representation of captions

i	Image feature vector	Partial caption	Target caption
1	Image_feature_vector	[6]	5
2	Image_feature_vector	[6, 5]	4
3	Image_feature_vector	[6, 5, 4]	3
4	Image_feature_vector	[6, 5, 4, 3]	8
5	Image_feature_vector	[6, 5, 4, 3, 8]	1
6	Image_feature_vector	[6, 5, 4, 3, 8, 1]	9
7	Image_feature_vector	[6, 5, 4, 3, 8, 1, 9]	3

4 Proposed Methodology

The model presented in [2] forms the basis for the proposed model. The proposed model is an attempt to see how an ensemble model of transformer encoder model [32] and bidirectional LSTM model [1, 33] will perform in comparison with models using just LSTMs or transformer encodings. In Figs. 2 and 3, the architecture of the transformer encodings model and BiLSTM is shown. The proposed model is an ensemble model of the transformer encodings model and BiLSTM in which the image acts an input and the output is a caption describing the image.

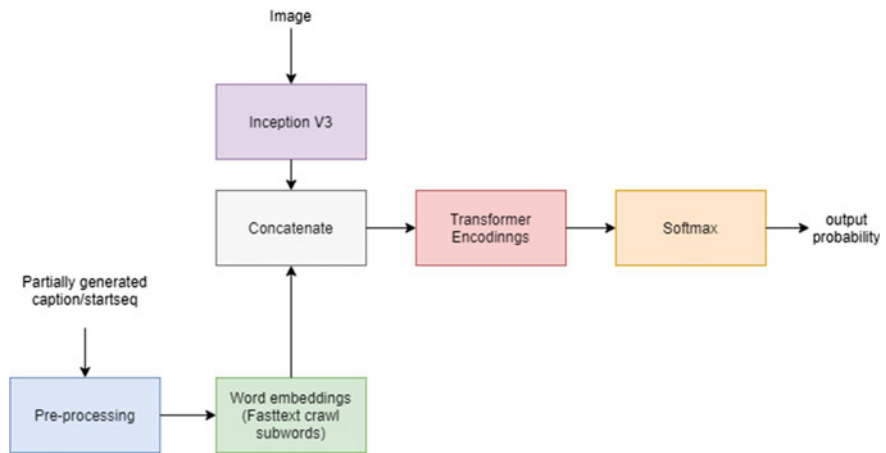


Fig. 2 Transformer encodings model architecture

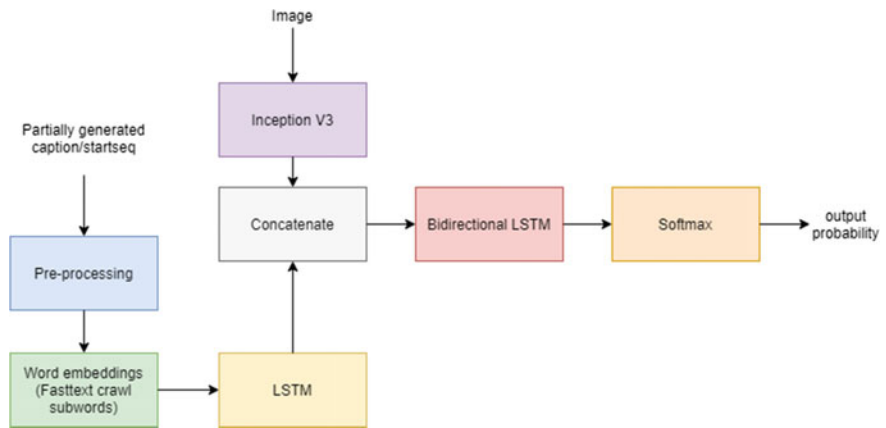


Fig. 3 BiLSTM model architecture

Before passing the partially generated caption or the startseq to the decoder, it is first passed through word embeddings. For the proposed model, FastText subwords crawl pre trained embeddings [34, 35] have been used. The encoder, decoder and the process of caption generation have been explained below.

#### ***4.1 Image Encoder***

The image encoder takes the image for which the caption is to be generated as the input and extracts the features relevant to predict the caption. For the proposed model, Inception-v3 [10] has been used for extracting features from the images. CNNs have been used extensively for the process of feature extraction. Inception-v3 is a pre-trained image classification model which contains multiple CNNs and is 48 layers deep. It has been trained on the ImageNet dataset which contains more than a million images. The pre-trained network can classify images into 1000 categories, but for the purpose of the proposed model, this is not required which is why the last layer of the inception model is replaced with a single fully connected layer.

#### ***4.2 Image Decoder***

The decoder portion of the model consists of the transformer encoders and a bidirectional LSTM. The model takes as input the output from word embeddings, and this is then passed to the transformer encoder. The transformer encoder model and the bidirectional LSTM model both calculate the individual probabilities of the next word of the caption. The two individuals probabilities are then combined by averaging these probabilities to output the next word of the caption.

#### ***4.3 Generation of Caption***

For the generation of the caption, the image along with the start token is passed to the model. This predicts the first word. In the next step, the features from the image along with the start token and first word are passed to the encoder to generate the second word. The same process is repeated till the end token is encountered or the max length has been reached. A max length of 25 has been used as most captions in the training data have a length less than this number. The sentence produced on completion of this process is the caption generated for the given image.

## 5 Results and Evaluation

As shown in the table above, the proposed methodology ensemble model shows more promising results for the first three images as compared to the existing models based either on LSTM or transformer. All models were trained on the Flickr30k dataset for 15 epochs with a learning rate of  $1e-4$ . The models also have a dropout of 0.5 after the concatenate layer. Apart from that, FastText crawl subword pre-trained word embeddings have been used to train the models. The LSTM model achieved a BLEU score [11] of 0.5913 and a loss of 0.3149 on a test set consisting of 500 unseen images. The transformer encodings model achieved a BLEU score of 0.5747 and a loss of 0.321 on a test set consisting of 500 unseen images. The ensemble model of LSTM and transformer encodings achieved that a BLEU score of 0.6192 shows a slight improvement over the existing models.

The first three images in Table 2 are clearly well defined and have unambiguous objects or features, and the Flickr30k dataset used to train all the models have similar distinct images with appropriate caption. Hence, the captions generated by the proposed model for the first three images scored better than the existing models. Whereas, for the remaining three images in Table 2, the features of the images are vague, in the background as clusters and not clear in the foreground or not well lit overall. Therefore, the scores of the proposed model or existing models are sub-par.

## 6 Conclusion






From our findings, we can conclude that the suggested ensemble model is superior at generating captions for unknown images than the existing models. The images can be categorised as follows:

- Images with human activities.
- Images of animals.
- Images of scenery.

To create a well-rounded model which produces good captions for each category, we need sufficient data for each category. The paper also concludes that training the proposed model with a large dataset will teach it to visualise the image better; for instance, if the training dataset included images having large clusters or more focus on the background instead of the foreground or dim lighting, then the proposed model would accordingly give better results for such images and communicate the information acquired by it to the user. The paper also shows that with the deep learning techniques available currently there is still scope for betterment; however, the current limitations of generating high scoring captions of distinct images and not clustered or slightly vague images need to be overcome.




**Table 2** Results on three clear images and three images with vague features

Image	Transformer + LSTM model Bleu score = 0.6192	Transformer model BLEU score = 0.5747	LSTM model BLEU score = 0.5913
	<p>Caption: a man in a blue shirt is standing on a beach BLEU score = 0.909</p>	<p>Caption: a man in a blue shirt is standing in front of a lake BLEU score = 0.593</p>	<p>Caption: a man in a blue shirt and black pants is standing on a beach BLEU score = 0.731</p>
	<p>Caption: a young boy is sliding down a slide Bleu score = 0.875</p>	<p>Caption: a young boy in a blue shirt is playing with a toy Bleu score = 0.666</p>	<p>Caption: a little girl in a pink shirt is playing with a toy Bleu score = 0.75</p>
	<p>Caption: two men playing soccer in a field BLEU score = 1.0</p>	<p>Caption: a young boy in a blue shirt is playing with a soccer ball BLEU score = 0.384</p>	<p>Caption: two soccer players are playing soccer BLEU score = 0.833</p>
	<p>Caption: a man is sitting on a bench in front of a large amount of people BLEU score = 0.266</p>	<p>Caption: a man in a black jacket is walking through a crowded area BLEU score = 0.416</p>	<p>Caption: a man is standing in a crowded area with a large amount of people in the background BLEU score = 0.352</p>
	<p>Caption: a group of people are walking down a street BLEU score = 0.555</p>	<p>Caption: a group of people are walking down a street BLEU score = 0.555</p>	<p>Caption: a man is standing on a city street in front of a large white building BLEU score = 0.4</p>

(continued)

Table 2 (continued)

Image	Transformer + LSTM model Bleu score = 0.6192	Transformer model BLEU score = 0.5747	LSTM model BLEU score = 0.5913
	Caption: a group of people are standing in front of a large building BLEU score = 0.416	Caption: a group of people are walking down a street BLEU score = 0.555	Caption: a group of men in orange uniforms are standing in front of a building BLEU score = 0.285

## References

1. Shah P, Bakrola V, Pati S (2017) Image captioning using deep neural architecture. In: 2017 international conference on innovations in information embedded and communication systems (ICIIECS)
2. Vinyals O, Toshev A, Bengio S, Erhan D (2016) Show and tell: lessons learned from the 2015 mscoco image captioning challenge. *IEEE Trans Pattern Anal Mach Intell* 99:1–1
3. Farhadi A, Hejrati M, Sadeghi MA, Young P, Rashtchian C, Hockenmaier J, Forsyth D (2010) Every picture tells a story: generating sentences from images. *European conference on computer vision*. Springer, Berlin, Heidelberg, pp 15–29
4. Yang Y, Teo C, Daumé III H, Aloimonos Y (2011) Corpus-guided sentence generation of natural images. In: *Proceedings of the 2011 conference on empirical methods in natural language processing*, pp 444–454
5. Kulkarni G, Premraj V, Ordóñez V, Dhar S, Li S, Choi Y, Berg AC, Berg TL (2013) Babytalk: understanding and generating simple image descriptions. *IEEE Trans Pattern Anal Mach Intell* 35(12):2891–2903
6. Mitchell M, Dodge J, Goyal A, Yamaguchi K, Stratos K, Han X, Mensch A, Berg A, Berg T, Daumé III H (2012) Midge: generating image descriptions from computer vision detections. In: *Proceedings of the 13th conference of the european chapter of the association for computational linguistics*, pp 747–756
7. Ordóñez V, Kulkarni G and Berg TL (2011) Im2text: describing images using 1 million captioned photographs. *Adv Neural Inf Proces Syst* 1143–1151
8. Karpathy A, Fei-Fei L (2015) Deep visual-semantic alignments for generating image descriptions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 3128–3137
9. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, vol 28, pp 1–9
10. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision. *arXiv preprint [arXiv:1512.00567](https://arxiv.org/abs/1512.00567)*, 37
11. Papineni K, Roukos S, Ward T, Zhu W-J (2002) Bleu: a method for automatic evaluation of machine translation. In: *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics, pp 311–318
12. Amritkar C, Jabade V (2018) Image caption generation using deep learning technique. In: 2018 fourth international conference on computing communication control and automation (ICCUBEA)
13. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)*
14. Hrga I, Ivašić-Kos M (2019) Deep image captioning: an overview. In: *MIPRO 2019*, Opatija Croatia
15. Fang H, Gupta S, Iandola F, Srivastava RK, Deng L, Dollár P, Gao J, He X, Mitchell M, Platt JC, Lawrence Zitnick C (2015) From captions to visual concepts and back. In: *CVPR*, pp 1473–1482
16. Ivašić-Kos M, Pobar M, Ribarić S (2015) Automatic image annotation refinement using fuzzy inference algorithms. In: *IFSA- EUSFLAT 2015*. Gijón, Asturias, (Spain), p 242
17. Vinyals O, Toshev A, Bengio S, Erhan D (2015) Show and tell: a neural image caption generator. In: *CVPR*, pp 3156–3164
18. Hodosh M, Young P, Hockenmaier J (2013) Framing image description as a ranking task: data, models and evaluation metrics. *JAIR* 47:853–899
19. Mathews A, Xie L, He X (2016) Senticap: generating image descriptions with sentiments. In: *13th AAAI conference on artificial intelligence*
20. Gan C, Gan Z, He X, Gao J, Deng L (2017) Stylenet: generating attractive visual captions with styles. In: *CVPR*, pp 3137–3146

21. You Q, Jin H, Luo J (2018) Image captioning at will: a versatile scheme for effectively injecting sentiments into image descriptions. arXiv preprint [arXiv:1801.10121](https://arxiv.org/abs/1801.10121)
22. Chen T, Zhang Z, You Q, Fang C, Wang Z, Jin H, Luo J (2018) Factual or emotional: stylized image captioning with adaptive learning and attention. In: ECCV, pp 519–535
23. Nezami OM, Dras M, Anderson P, Hamey L (2018) Face-cap: image captioning using facial expression analysis. In: Joint European conference on machine learning and knowledge discovery in databases, pp 226–240
24. Rashtchian C, Young P, Hodosh M, Hockenmaier J (2010) Collecting image annotations using Amazon’s Mechanical Turk. In: Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon’s Mechanical Turk, pp 139–147
25. Young P, Lai A, Hodosh M, Hockenmaier J (2014) From image descriptions to visual denotations: New similarity metrics for semantic inference over event descriptions. *Trans Assoc Comput Linguist* 2:67–78
26. Chen X, Fang H, Lin TY, Vedantam R, Gupta S, Dollár P, and Zitnick CL (2015) Microsoft COCO captions: data collection and evaluation server. arXiv preprint [arXiv:1504.00325](https://arxiv.org/abs/1504.00325)
27. Papineni K, Roukos S, Ward T, Zhu WJ (2002) BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting on association for computational linguistics, pp 311–318
28. Denkowski M, Lavie A (2014) Meteor universal: language-specific translation evaluation for any target language. In 9th workshop on statistical machine translation, pp 376–380
29. Lin CY (2004) Rouge: a package for automatic evaluation of summaries. In Text summarization branches out: proceedings of the ACL-04 workshop, vol 8
30. Vedantam R, Zitnick C, Parikh D (2015) Cider: consensus based image description evaluation. In: CVPR, pp 4566–4575
31. Anderson P, Fernando B, Johnson M, Gould S (2016) Spice: semantic propositional image caption evaluation. In: ECCV 2016, pp 382–398
32. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, Kaiser L, Polosukhin I (2017) Attention is all you need
33. Zhou C, Sun C, Liu Z, Lau F (2015) A c-lstm neural network for text classification
34. Joulin A, Grave E, Bojanowski P, Douze M, Jégou H, Mikolov T (2016) Fasttext.zip: compressing text classification models. In: CoRR, abs/1612.03651
35. Bojanowski P, Grave E, Joulin A, Mikolov T (2016) Enriching word vectors with subword information. In: CoRR, abs/1607.04606

# Graph-Based Transductive Label Propagation Approach for Multi-label Classification



S. C. Dharmadhikari, S. H. Chandak, and R. R. Chhajed

**Abstract** It is evident that conventional approaches for multi-label text classification are striving for several labeled instances in order to train classifier, and they are often unable to utilize unlabeled data. The demand of many applications is not only inclining toward automated set up wherein need for the manual assignment of multiple labels to each document will be removed but there is also more reliability in the label set allocation process. In an attempt, graph-based label propagation approach in multi-label setting for specifically text classification application is presented in this paper. The approach is scalable and mainly based on concept of transductive learning. It is holistic in the sense that it is having capacity to effectively utilize unlabeled data, extraction of relevant features from huge feature set, cope up with exponential growth in document collection, labels and exploration of correlativity among the labels for classification. Performance of the proposed system is demonstrated by means of comparison with standard systems on classical datasets involving multi-labeled data.

**Keywords** Multi-label text classification · Information retrieval · Transductive learning

## 1 Introduction

This paper addresses various issues which are prevalent in multi-label text classification system. Traditional classification problems are associated with learning and classifying documents into single classes from given set of classes. When only two classes are considered, then it is referred to as a binary classification problem and when there are more than two classes in the set, it can be termed as multi-class problem. But, both problems only assign one class to an example. However, in multi-label case, each instance may belong to multiple classes simultaneously

---

S. C. Dharmadhikari (✉) · S. H. Chandak · R. R. Chhajed  
Department of IT, PICT, Pune, India

and thus more complex to implement as well as more demanding in realistic classification problems. This requirement makes it more challenging. Unfortunately, in spite of having more promising application area, multi-label problems have not been investigated as extensively as other counter parts.

The major objective of any multi-label text classification system is to determine the correct set of classes for document instance under consideration based on some experience (previously trained samples). For most of the applications, traditional supervised learning set up is used to implement multi-label text classifiers. But, dependency on availability of label set, and otherwise manual labeling time makes it infeasible in many cases. Moreover, it has been observed in many real-world scenarios that there is existence of abundant unlabeled data. Hence, in our framework, we considered transductive set up to effectively utilize unlabeled data along with available labeled data.

With this insight, we presented well known state-of-art in Section 2 of this paper. Whereas Section 3 presents the methodology and algorithm behind the transductive framework. Subsequently, Sect. 4 presents experimental set up and result discussion.

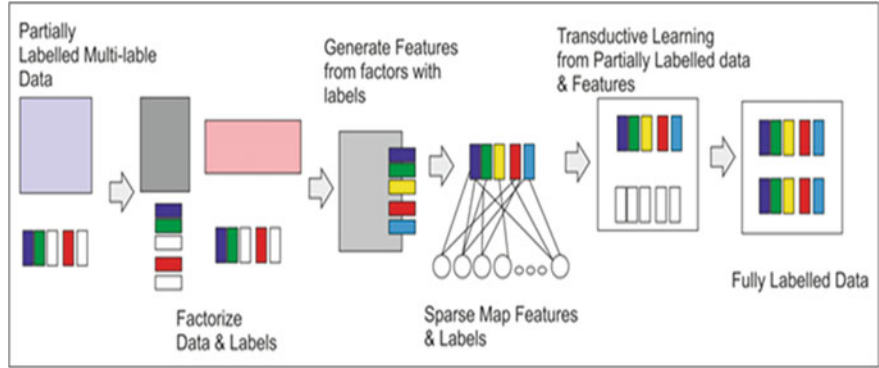
## 2 Literature Survey

Multi-label classification methods have been investigated extensively and generally belong to any of the type of group as problem transformation (PT) methods, algorithm adaptation (AA) methods, ensemble based methods (EB), large scale data processing methods (LSD), shared subspace based methods (SSB), input encoding based methods (IEB), online multi-label methods (OM), generative modeling based methods (GM), hierarchical based methods (HB), correlation based methods (CB), label transformation based methods (LT), label subset selection based method (LSS), semi-supervised learning based methods (SSL) as given in [10]. Table 1 summed up major algorithms investigated so far under their category as.

It has been observed that algorithms belonging to every group is addressing some unique features, but none of the methods uniformly addressing the issue of utilization of unlabeled data along with labeled data, label correlation, utilization of relationship between objects and class label set for classification and large-scale data [10]. Semi-supervised learning based methods are utilizing most of the available unlabeled data along with labeled data efficiently but not considering aforesaid three remaining issues. The consideration of all these factors may be useful for more performance improvement of these methods. This fact highlights the need for investigating the more holistic approach with capacity to consider the intra label, intra text and inter text-label relationship as well [10].

**Table 1** Summed up major algorithms investigated so far under their category as

Multi-label classification methods	Examples of Algorithms
PT	BR, LP Based, PW Based: RPC, CLR, QWML; CC,PS, PPT
AA	ML-C4.5, SVM-HF, AdaBoost.MH, AdaBoost. MR BP-MLL, ML-kNN
EB	RDT, RAKEL, EPC, EPS, ECC, RPC
LSD	MetaLabeler
SSB	MSSBoost
IEB	INSDIFF
OM	Online MLAL, BOMC
GM	PMM CoLModel, CRF, CML, CMLF
HB	HBR, HOMER
CB	PCC, EPCC
LT	PLST, CPLST, CCA
LSS	MOPLMS, ML-CSSP
SSL	CNMF, SSNMF, Graph based, BR + ,SSBR, SSLP, SSRakEL



**Fig. 1** Graph-based transductive label propagation approach for multi-label text classification

3 Methodology

Figure 1 shows the detailed architecture of the proposed approach referred to as GTLP (graph-based transductive label propagation) and has three main components viz. Multi-label feature extraction, sparcification map preparation for labels and data for balanced matching for concurrence between text and labels, transductive learning from partially labeled data and features for enhancing label assignment process to unlabeled data set by verifying eligibility of labels. The framework is holistic in the

sense it covers three main paradigms of classification system namely preprocessing, feature selection and extraction, classification.

Text collection is served as input to the setting in the initial testing phase. This collection is subsequently transformed into the feature set and stored in the .csv format. It consists of partially labeled data as well as unlabeled data. Multi-label feature extraction part of the framework aims mainly at transforming the entire text collection into feature vector form and reduces it further by retaining most relevant ones. It is not only applying this approach on text features but also on label set which is represented as a feature set in order to yield more reliable outcome. Then, these text and label features are transformed into data and label matrix form. Label matrix is constituted by representing every class label as an equivalent binary vector form. Furthermore, non-negative matrix factorization in multi-label setting is applied in order to decompose the data and label matrix and thereby reduced feature set in the matrix form is getting produced.

Subsequently, correlativity and closeness factor between intra text features, intra label features, inter text and label features are explored by constructing sparse feature map using balanced matching method. Most, relevant pairs are selected for subsequent training and classification phase. Thereafter, classification is carried out by adapting transductive learning process of semi-supervised learning wherein mainly available labeled data set is firstly utilized to label unlabeled data. We have added our coating of eligibility checking procedure which will finalize the assignment of label set to unlabeled data only if it crosses the set threshold, thereby increasing reliability of assignment of label set. The representative algorithm is given below:



```

/* Transductive Learning on Multi-Label Data */

Input: Document Collection D, Label Set L
Output: Label Set Assignment

/* Preprocess Document Collection and prepare data matrix*/

D_t = tokenize(D) #Tokenize all documents
D_clean = stop_word_removal(D_t) # Remove stop words
A = TFIDF_Transformer(D_clean) # Generate TFIDF Model
B = OneHotEncode(Label)

W,H = Factorize(A)
J,I = Factorize(B)

W    = B_Matching(sparsification(ACOS),ACOS)
W_L = B_Matching (sparsification(BCOS),BCOS)

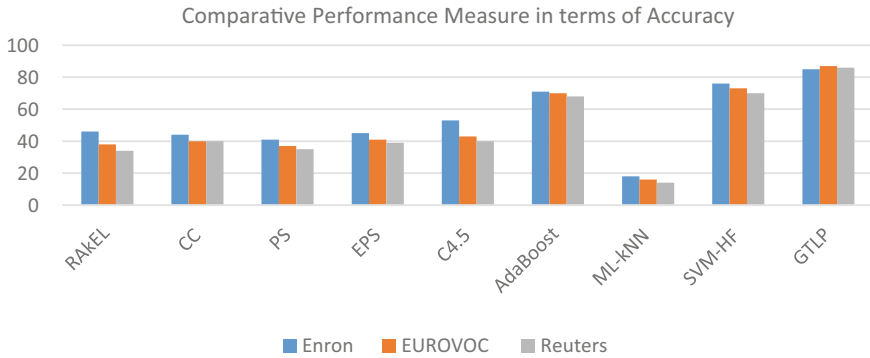
// Train the classifier using labeled samples
// Predicting label set and adding its Eligible labels in the training set
// Obtain Initial entries of Final Label Information Matrix F
F = W_L

Repeat until Convergence
{
//Propagate labels of labeled data towards unlabeled da-ta
Tmp_F = FT * W_L[1] ;
// Verify Eligibility of propagated labels
for each entry in sparse feature vector matrix
    for each entry in sparse label vector matrix
        Assign label set to unlabeled data
    Check for Eligibility of new label set
    If Eligible propagate it for inclusion
}

```

## 4 Experimental Steps and Results

The experiment had been performed on three datasets of Enron, Eurovoc and Reuters (RCV1) in multi-label setting in order to measure the performance of GTLP. Experimental set up is built using MEKA and MULAN libraries. Datasets are chosen in such a way that their cardinality varies from small scale to very large scale and they also represents different type of underlying data from Email information, legal information to newswire stories. The accuracy is measured against these datasets for GTLP. Then, in order to verify where the performance of GTLP stands against



**Fig. 2** Comparison of accuracy using various supervised techniques and GTLP

supervised counterpart, empirical analysis being carried out on existing multi-label approaches namely RAkEL, CC, PS, EPS, C4.5, AdaBoost, ML-kNN, SVM-HF. The comparative performance in terms of bar graph is shown in Fig. 2.

It is evident from Fig. 2 that using GTLP overall classification accuracy of multi-label text classification has been increased substantially by on an average 18–20% as compared to others. The consideration of utilization of unlabeled data, extraction of relevant features from huge feature set, scaling capacity to cope up with exponential growth in document collection, labels and exploration of correlativity among the labels for classification makes the GTLP more robust and showing outperformance as compared to others.

## 5 Conclusion

This paper presents the attempts taken to improve overall scope and learn ability of multi-label text classifier by means of employing transductive learning based set up. The holistic approach here is to modify learning system at multiple levels. The framework is holistic in the sense it covers three main paradigms of classification system namely preprocessing, feature selection and extraction, classification. Novel feature extraction technique for multi-label setting is applied here. At next level, b-matching based sparse graphs address complexity arises due to increased number of document and label spaces. Furthermore, a unique approach is devised for inclusion of reliable training instances by utilizing unlabeled data by extending functionality of traditional label propagation-based approaches. However, in future inclusion of hybrid text representation scheme and incremental learning approach may be also serve the game changer for attempts in performance improvements.

## References

1. Liu W, Wang J, Chang S-F (2012) Robust and scalable graph-based Semi supervised learning. In: Proceedings of the IEEE, vol 100, no 9, Sept 2012
2. Nigam K (2000) Semi-supervised text classification using EM. *J Mach Learn* 39(3):103–134
3. Zha Z-J et al (2009) Graph-based semi-supervised learning with multi-label. *J Visual Commun Image Representation* 20:97–103
4. Liu Y et al (2006) Semi-supervised multi-label learning by constrained non-negative matrix factorization. In: Proceedings of artificial intelligence in association for the advancement of artificial intelligence conference, vol 1, pp 421–426
5. Chen G et al (2008) Semi-supervised multi-label learning by solving a Sylvester equation. In: Proceedings of the 2008 SIAM international conference on data mining, Atlanta, Georgia, pp 410–419
6. Lee H et al (2010) Semi-Supervised non-negative matrix factorization. In: *IEEE signal processing letters*, vol 17, pp 4–7
7. Wei Q et al (2009) Semi-supervised multi-label learning algorithm using dependency among labels. In: International conference on machine learning and computing, pp 112–116
8. Rodrigues SC (2009) Using confidence value in multi-label classification problems with semi supervised learning. In: Proceedings of the LNCS international conference on machine learning, pp 126–132
9. Cherman et al (2012) Incorporating label dependency into the binary relevance framework for multi-label classification. *J Expert Syst Appl* 39:1647–1655
10. Dharmadhikari S, Ingle M (2016) Intelligent semi supervised multi label text classification. Ph.D dissertation, School of Computer Science & Information Technology, Devi Ahilya Vishwavidyalaya, Indore. Accessed on: May 4, 2021 (Online). Available: <http://hdl.handle.net/10603/111775>

# Detection and Classification of Diseases and Maturity of Dragon Fruits



Vrushali Kulkarni, Pranali Kosamkar, Chanavi Singh, Purva Ingle, and Vedant Modi

**Abstract** A huge population of India is dependent on agriculture as their primary source of livelihood. The financial growth of India depends on the GDP (gross domestic product) rate of the agricultural commodities. Since dragon fruit is in high demand both domestically and globally, dragon fruit plantations are economical for small scale to large scale growers and businesses. There is a heavy focus on each stage of the fruit's development because prolonged or premature harvesting of the fruit can result in food decay and can also impact the India's economic growth. Farmers are still making use of manual methods of grading fruit maturity and identifying disease ridden crops, which can result in poor grading due to weariness and misjudgment. The article proposes the use of region-based convolutional neural network, a deep learning algorithm for detection and classification of common dragon fruit diseases and detection of dragon fruit maturity. There are three key steps to this process of detection and classification. The first step includes identification of several regions of interest, which are the bounding box candidates using selective search. A broad convolutional neural network extracts features from each region in the second step. The last step performs grouping of support vector machine (SVM) that tries to classify the regions for the classification of different dragon fruit diseases and detecting its maturity.

**Keywords** Region-based convolutional neural network · Image processing · Image classification · Agriculture · Fruit diseases

## 1 Introduction

Dragon fruit (*Hylocereus* spp.), a perennial herbaceous climbing cactus commonly known as Red Pitaya, has recently aroused the interest of Indian growers, not only

---

V. Kulkarni (✉) · P. Kosamkar · C. Singh · P. Ingle · V. Modi  
MIT World Peace University, Pune, India  
e-mail: [vrushali.kulkarni@mitwpu.edu.in](mailto:vrushali.kulkarni@mitwpu.edu.in)

P. Kosamkar  
e-mail: [pranali.kosamkar@mitwpu.edu.in](mailto:pranali.kosamkar@mitwpu.edu.in)

for its attractive red or pink hue and economic value as a fruit, but also for its high-antioxidant capacity, vitamins, and minerals content [1]. It is one promising crop that can be easily grown in degraded land and drought-prone areas across the world. Dragon fruit processing could be a profitable avocation for both backyard growers and entrepreneurs of average and large-scale plantations, thanks to the major requirement in both domestic and foreign markets [1]. It has great potential for quick returns on investment. It is a source of fast stable income, hence making it popular among farmers. Dragon fruit cultivation has already begun in various parts of India, with numerous farmer success stories from various regions. In 2020, India's dragon fruit production increased dramatically to over 12,000 MT over a 3000–4000 ha area. These projections are based on first-hand information gathered by the ICAR–NIASM from progressive growers, entrepreneurs, consultants, and state agricultural department officials across the country. More than 80% of the total 3085 ha (2468 ha) is under new cultivation with a plantation age of less than 18 months [2].

Computer vision-based systems have been widely used in the agricultural and food organizations, particularly in the areas of quality control and product classification. In the fruit-based industries, quality management and marketing are critical for producing high-quality goods. Dragon fruit quality control has traditionally been done by humans. The different types of deformities in form and shape make dragon fruit grading difficult.

In this paper, we propose to use region based convolutional neural network (R-CNN) to classify and detect diseases that plague the dragon fruit. Anthracnose, brown spots, and stem rots are common diseases that affect dragon fruit crops (fungal and bacterial pathogens) [1]. We also aim to detect fruit maturity. R-CNN uses mechanisms to extract the region of interest. Post extraction the regions are fed into a regular convolutional neural networks (CNN). CNN has been widely used in the agricultural domain for fruit maturity and disease detection because of its high accuracy. In R-CNN, high-capacity CNNs can be used to localize and segment artifacts from bottom-up region proposals and when labeled training data is limited, supervised pre-training for an auxiliary task, followed by domain-specific fine-tuning yields a major improvement in performance [3].

## 2 Literature Survey

Various methods for detecting and classifying fruit diseases have been proposed and developed. The appearance of surface defects has an effect on the quality and worth of fruits and vegetables, and removing fruits and vegetables with severe defects early will help avoid contamination of the whole patch. As a result, the quality grading of fruits and vegetables is the general application of image analysis. Traditional computer vision inspection of fruits and vegetables for color, texture, scale, and form is now automatic in industrial sorting machines. Under CNN, many different models have produced better accuracies and more optimized models given technological constraints.

Alharbi et al. have used five models of CNN for classification of healthy apples or apples ridden with diseases. The dataset contains a total of 3200 images. Different ratios have been considered for dividing the dataset into training and testing sets, where every ratio is trained and tested ten times. Model-5 gave us the best accuracy which is 99.17% where the training dataset was 90% and the testing dataset was 10% of the total dataset [4]. Zhao et al. selected YOLOv2 as the experimental model and applied the improved YOLOv2 method for the detection of healthy and diseased tomatoes. The dataset consists of images belonging to 7 classes. Four image processing methods were applied on the dataset of 200 images, which yielded 800 augmented images. K-means clustering algorithm was used in the training set, so as to make improvements in the model performance. The mean average precision of this network is 97.24% [5]. Singh et al. explore different diseases and infections in coconut trees using deep learning algorithms. A custom CNN model was trained and pre-trained CNN models InceptionV3, VGG16, DenseNet201, VGG19, MobileNet, InceptionResNetV2, Xception, and NASNetMobile were fine tuned. These algorithms were applied on a dataset of 1564 images. Among the pre-trained CNNs, MobileNet had the greatest accuracy of 82.10%. Moreover, the CNN achieved an accuracy of 85.63% and 96.94% with two and four convolutional layers, respectively [6]. Azizah et al. make use of CNNs to detect defects on the surface of the mangosteen fruit. A four-cross validation on 120 test images resulted in a 97.5% accuracy. A CNN model with 50 epochs with filter width 2 and pool size 2 and stride 2 gave the best results [7]. Singh et al. suggest a multilayer convolutional neural network (M-CNN) to classify mango diseased leaves. Their dataset consists of a total of 2200 images categorized into four classes. Tensorflow framework has been used here. The model had an accuracy of 97.13% [8]. Kukreja et al. a dense CNN algorithm is used to detect the defects of citrus fruit. The dataset contains a total of 1200 images. In order to improve the performance of the dense CNN model, the proposed model makes use of data augmentation and data preprocessing. The accuracy of the proposed model is 89.1% [9].

Nagasubramanian et al. in their work use hyperspectral imaging to create 3D images of the soybean crops to detect the charcoal rot disease using deep CNNs. They achieved a classification accuracy of 95.73%. The aim of the study was to prove the efficiency of deep learning methods in performing such tasks [10]. Saleem et al. review the detection and classification of plant diseases using deep learning models. They conclude that deep learning models effectively perform the desired tasks in different scenarios. They emphasize the need for early detection of diseases and describe how deep learning is successful in doing that. Alex-Net is one of the most popular CNN models used [11]. Raikar et al. classify and grade okra on certain parameters such as shape, size, and color. AlexNet, GoogLeNet, and ResNet50 CNNs are used for this purpose on a dataset of 3200 images. ResNet50 took the longest training time, and AlexNet the least. ResNet50 was found to achieve an overall accuracy of 99.17%, GoogLeNet of 68.99%, and AlexNet, the least of 63.45%. However, the tenderness, freshness and infected portions of the vegetable sample cannot be identified using the given models [12].

Ozguven et al. propose a method to automatically detect and classify leaf spot diseases in sugar beet. An updated faster R-CNN algorithm and a faster R-CNN algorithm is used on a dataset of 155 images. The faster R-CNN classified some healthy areas as diseased and couldn't detect certain diseased areas. The updated model proved to be better at detection of diseased areas with an accuracy of 95.48% [13]. Parvathi et al. detect the maturity stages of coconuts in complex backgrounds using faster R-CNN. This model with ResNet50 network was compared with SSD, YOLOv3, and R-FCN models. Although the YOLOv3 model had the shortest detection time per image, the faster R-CNN model with ResNet50 was found to be the most suitable for the task [14]. Wang et al. proposed an improved mask-RCNN based model for detection of lesion defects on fruits. The dataset consists of 500 images of each, apples, peaches, pears, and oranges with spots on their surface. The mean average precision of the algorithm in detection of lesions is greater than 95%. The proposed model has high accuracy but the speed of the detection process is slightly slower [15]. Ni et al. aim to count blueberry berries in a cluster and assess their maturity using deep learning methods. They make use of around 724 outdoor images of blueberry clusters and use mask R-CNN to measure classification and segmentation performance. They have considered 3 traits namely, number, maturity, and compactness. Their mask accuracy is 90.6% for a precision of 78.3%. Since it considers 2-D images, the count of the berries in a cluster can be incorrect at times, since the berries hidden by the berries in front would not be detected [16].

Despite the extensive use of deep learning techniques for detecting diseases and fruit maturity in a variety of fruits, not a lot of work has been done on dragon fruits. Khisanudin et al. proposed to classify dragon fruits based on their maturity using image processing. Their aim was to reduce human intervention by using Naive Bayes algorithm on HSV color space. With a dataset of 150 images, they achieved an accuracy of 86.6% [17]. Another paper by Dong et al. proposes to detect diseases on the surface of dragon fruit using image segmentation. They make use of the FCM algorithm and two-dimensional optimized OTSU algorithm to achieve background and foreground segmentation [18]. Wismadi et al. make use of VGGNet like network to detect the ripeness of dragon fruits. They achieve an accuracy of 91% [19].

## **2.1 Key Findings**

Based on the papers, we surveyed pertaining to detection and classification of diseases, and predicting the maturity level of agricultural products, we found some limitations.

Firstly, the most difficult issue with deep learning tasks is a lack of effective training samples with labels. To the best of our knowledge, we did not find a freely available and a large enough labeled dataset of dragon fruits. Frequently, datasets include a large number of photographs of various plant species with diseases against a simple backdrop. A real-world background with complexities should be considered for a realistic scenario. Moreover, some algorithms fail to segment relevant features

required for the classification of certain classes. The accuracy can be improved by dealing with the problems in the images of the real-time database due to variation in temperature, shadowing, overlapping of leaves, and the presence of multiple objects. Additionally, several images in datasets, often contain shadows, overlapping of leaves, and complex backgrounds which makes the task of the algorithm more challenging. This may result in lower accuracies. Another important factor to consider is the model's ability to detect diseases in crops at different stages of the disease growth.

### 3 Proposed Methodology

Convolutional neural network as a winning model of deep learning has shown exceptional performance in various applications such as object detection, image analysis, and image classification. Our proposed work utilizes a deep learning model namely R-CNN for the detection of ripeness of dragon fruit and the detection and classification of important diseases that affect dragon fruits viz. anthracnose, brown spots, and stem rot.

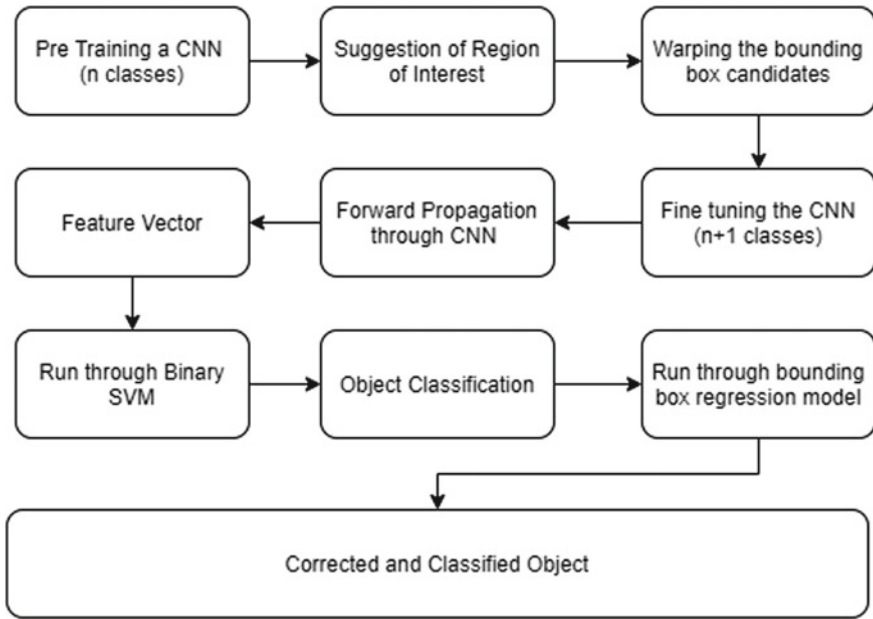
Application of high-capacity convolutional neural networks (CNN) to bottom-up region proposals resulted in the model we now call region-based convolutional neural networks. Using this model, we can localize objects and segment them. When the labeled training dataset is sparse, supervised pre-training for an auxiliary task, and subsequently domain specific fine-tuning bears a notable augment in the performance. The method consists of three main steps. In the first step, it identifies a number of regions of interest, which are the bounding box candidates [3]. It creates bounding boxes using this process which is called selective search. We use this because, it considers multiple grouping criteria like color, texture, and brightness [20]. In the second step, a large convolutional neural network extracts features from each region. The third step makes use of SVM classes for classification of the regions [3].

#### 3.1 Proposed Workflow

##### Steps Involved.

1. Pre training a CNN network, for classification of 'n' classes.
2. Using selective search suggestions of category independent regions of interest or bounding box candidates of varying sizes that may contain our objects [20].
3. To achieve a fixed size needed by CNN, warping the bounding box candidates.
4. Fine tuning the CNN on the warped regions for ' $n + 1$ ' classes (+1, since we are including the background).





**Fig. 1** Proposed workflow

5. Forward propagation through the CNN for each image region gives a feature vector which is then run through a binary SVM, independently trained for different classes.
6. The box is then run through a regression model in order to make corrections to the box after object classification has been done [21] (Fig. 1).

### 3.2 Image Processing

**Feature Selection** In order to detect the maturity of the fruit as well as lesions caused due to aforementioned diseases, the color statistical features of the R (Red), G (Green), B (Blue) values of the fruit. We also take into consideration the shape of the fruit to help detect its ripeness.

*Color Statistical Features.* The color statistical features can be computed using the following formula [22]

$$\text{Color Mean}(\mu) = \frac{1}{N} \sum_{i=1}^N P_i \quad (1)$$

$$\text{Standard Deviation}(\sigma) = \left( \frac{1}{N-1} \sum_{i=1}^N (P_i - \mu)^2 \right)^{\frac{1}{2}} \quad (2)$$

$$\text{Skewness} = \frac{\sum_{i=1}^N (P_i - \mu)^3}{N\sigma^3} \quad (3)$$

**Shape Feature Extraction.** Specific usage of size-dependent features to assess the form of fruit and vegetable species with extremely non-uniform shapes is insufficient. More efficient features independent of size for shape definitions and measurements independent of size measurements include boundary encoding, invariant moments, and fourier descriptors. A chain code is used in boundary encoding to define the form [23].

**Defect Detection.** We use the defect ratio, defect medial axis length, and defect perimeter as basic geometric features. The defect ratio is the number of pixels in the defected region divided by the number of pixels in the fruit region [24].

### 3.3 Performance Metrics

We propose using the following measures to assess our model's performance: accuracy, precision, recall, F1 score, and a confusion matrix as True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN) [25].

**Accuracy.** The number of correct predictions made out of all predictions made by the model. When the data is symmetric, accuracy is a favorable determining factor, but when the data is imbalanced (positive and negative classes have a very different number of data instances), it fails to provide a clear picture [26].

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} \quad (4)$$

**Precision.** The number of positive predictions that are correct out of total positive predictions. A low false positive rate is indicated by a high precision. This tells us how often a model is correct when it predicts a positive outcome.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

**Recall.** The ratio of correct positive predictions out of all positive cases.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

**F1 Score.** This accounts for both precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

**Fig. 2** Confusion matrix  
(Source [25])

		Predicted Label	
		Negative	Positive
True Label	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

**Confusion Matrix.** This is a result of the number of correct and incorrect predictions totaled and broken down by class using count values (Fig. 2).

## 4 Conclusion

This article proposes an approach which incorporates R-CNN, a deep learning technique for the recognition and classification of dragon fruit diseases and detection of dragon fruit ripeness. Image processing is carried out to extract features based on three attributes namely color, shape and defect. There is a heavy focus on each stage of the fruit's development because prolonged or premature harvesting of the fruit can result in food deterioration and can also affect the India's economic growth. Farmers are still grading fruit maturity and recognizing infected crops using manual methods, which can lead to poor grading due to fatigue and error of judgment. The proposed computer vision and deep learning model help to enhance the process and make it more efficient. In future, we can use different deep learning models for same purpose.

## References

1. Karunakaran G, Arivalagan M, Sriram S (2019) Dragon fruit country report from India. ICAR-Indian Institute of Horticultural Research, Bengaluru, Karnataka, India
2. Wakchaure GC, Kumar S, Meena KK, Rane J, Pathak H (2020) Dragon fruit cultivation in India: scope, marketing, constraints and policy issues. Director, ICAR- National Institute of Abiotic Stress Management, Baramati, Pune, Maharashtra, India
3. Girshick R, Donahue J, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation [v5], pp 1–7

4. Alharbi AG, Arif M (2020) Detection and classification of apple diseases using convolutional neural networks. In: 2020 2nd International Conference on Computer and Information Sciences (ICCIS), pp 1–6
5. Zhao J, Qu J (2019) A detection method for tomato fruit common physiological diseases based on YOLOv2. In: 2019 10th international conference on Information Technology in Medicine and Education (ITME), pp 559–563
6. Singh P, Verma A, Alex JSR (2021) Disease and pest infection detection in coconut tree through deep learning techniques. In: Computers and electronics in agriculture, vol 182, 105986, ISSN 0168-1699
7. Azizah LM, Umayah SF, Riyadi S, Damarjati C, Utama NA (2017) Deep learning implementation using convolutional neural network in mangosteen surface defect detection. In: 2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp 242–246
8. Singh UP, Chouhan SS, Jain S, Jain S (2019) Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. IEEE Access 7:43721–43729
9. Kukreja V, Dhiman P (2020) A Deep Neural Network based disease detection scheme for Citrus fruits. In: 2020 International Conference on Smart Electronics and Communication (ICOSEC), pp 97–101
10. Nagasubramanian K, Jones S, Singh AK et al (2019) Plant disease identification using explainable 3D deep learning on hyperspectral images. Plant Methods 15:98
11. Saleem P, Arif M (2019) Plant disease detection and classification by deep learning. Plants 8:468
12. Raikar MM, Meena SM, Kuchanur C, Girraddi S, Benagi P (2020) Classification and grading of Okra-ladies finger using Deep Learning. In: Procedia computer science, vol 171, ISSN 1877-0509, pp 2380–2389
13. Ozguven MM, Adem K (2020) Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. In: Physica A: statistical mechanics and its applications, vol 535, 122537. ISSN 0378-4371
14. Parvathi S, Selvi ST (2021) Detection of maturity stages of coconuts in complex back-ground using Faster R-CNN model. Biosyst Eng 202:119–132. ISSN 1537-5110
15. Wang H, Mou Q, Yue Y, Zhao H (2020) Research on detection technology of various fruit disease spots based on mask R-CNN. In: 2020 IEEE International Conference on Mechatronics and Automation (ICMA), pp 1083–1087
16. Ni X, Li C, Jiang H et al (2020) Deep learning image segmentation and extraction of blueberry fruit traits associated with harvestability and yield. Hortic Res 7:110
17. Khisanudin S (2020) Murinto 2020. IOP Conf Ser: Mater Sci Eng 771:012022
18. Dong W, Xia Y, Liu Y (2020) Dragon fruit disease image segmentation based on FCM algorithm and two-dimensional OTSU algorithm. In: 2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), pp 969–973
19. Wismadi IM, Khrisne DC, Suyadnya IMA (2020) Detecting the ripeness of harvest-ready dragon fruit using smaller VGGNet-like network. J Electrical Electronics Informatics
20. Uijlings JRR, van de Sande KEA, Gevers T et al (2013) Selective search for object recognition. Int J Comput Vis 104:154–171
21. Object Detection for Dummies Part 3: R-CNN Family. <https://lilianweng.github.io/lil-log/2017/12/31/object-recognition-for-dummies-part-3.html>. Accessed on 29th April, 2021 at 10:00 pm
22. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syste Man Cybernetics SMC-3(6):610–621
23. Zhang B, Huang W, Li J, Zhao C, Fan S, Wu J, Liu C (2014) Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: a review. Food Res Int 62:326–343
24. Moallem P, Serajoddin A, Pourghassem H (2017) Computer vision-based apple grading for golden delicious apples based on surface features. Information Process Agriculture 4(1):33–40

25. Understanding Confusion Matrix. <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>. Accessed on 29th April, 2021 at 2:21pm
26. Baratloo A, Hosseini M, Negida A, El Ashal G (2015) Part 1: simple definition and calculation of accuracy, sensitivity and specificity. *Emergency (Tehran, Iran)* 3(2):48–49

# Implementation of DevOps Culture in Customizing Serverless Computing and Function as a Service (FaaS) for IoT Mashup



M. Ganeshan  and A. Rajesh 

**Abstract** The IoT is known as sprawling set of technologies. Cloud computing aids IoT and usage of cloud platform has helped in the rapid growth of IoT. The appearance of cloud platform serverless architectures, specifically function as a service (FaaS), gives rise to the question of the suitability of using them in IoT mashup environments. DevOps culture with a serverless architectural design with the addition of FaaS can effectively improve the computational capabilities that are closer to the end-user and provide data flow in a secure framework. In DevOps applications, FaaS has become a prime requirement for any cloud computing real-time data processing. It can be used to transform a single function code or unpacked to many micro-function codes. Apache OpenWhisk removes the chance of human application deployment errors compared to the previously manual evaluation, validation process and allows to customize serverless with many merits over AWS lambda service.

**Keywords** Serverless computing · Function as a service · DevOps culture · IoT mashup · AWS lambda · AWS IoT core · Serverless open source · Apache OpenWhisk · Kafka · Aws Greengrass

## 1 Introduction

Google states that the large count of IoT heterogeneous devices will exceed non-IoT by 2025. DevOps culture is in high demand in the field of IoT which is rapidly growing in the recent years [2]. In DevOps applications, FaaS has become an over most requirement for any cloud computing real-time IoT device message/data processing that allows to transform to a single function code from different IoT device rules and plays a prime role in automating micro-function codes which result in infrastructure deployment [1]. Serverless, in addition to FaaS architecture brings advantages to the

---

M. Ganeshan (✉) · A. Rajesh

Department of Computer Science and Engineering, Jain Deemed to be University, Bangalore, India

A. Rajesh

e-mail: [a.rajesh@jainuniversity.ac.in](mailto:a.rajesh@jainuniversity.ac.in)

DevOps process. It offers practical resource managements with transparency and freedom for developers to monitor most of the infrastructure components related to cloud application process. Any serverless computing method binds Dev and Ops teamwork with close proximity and even maintains marginal gap between Dev and Ops specialists [7]. But serverless services like lambda have some demerits like no control over the infrastructure environment since industry-defined tools and complex to call patterns are used. We aimed to create a prototype cloud architecture that explains the possible customization that can be performed on serverless computing features using open-source tools and the prototype architecture demonstrates to be more transparent, reduces complexity and is more applicable for custom-made IoT mashup architecture [6].

## 2 Background

### 2.1 *AWS IoT Core*

AWS IoT core service is used to connect AWS IoT Greengrass open-source runtime application outside the cloud environment and client IoT devices. These connections/communications can be established with cloud platform without the requirement of protocol design or manage server software for infrastructure maintains. IoT core can reinforce IoT devices for nearly billions of heterogeneous IoT mashup devices and trillions of messages sent by broker in terms of IoT rules. These rule messages/rules can be analyzed and routed to database for storage. Later, according to the requirement, the message can be streamed to endpoints designed for data analysis [4]. By using AWS IoT core, we can connect IoT devices to AWS platform, infrastructure and software services. We can monitor, track, interact with application process and communicate using device message data with all IoT mashup endpoints during run time. IoT devices can be connected to the cloud through Greengrass even in the absence of a network connection [4]. Greengrass provides benefits in automated configuration and authentication upon an IoT device first connectivity to AWS IoT core and provides end-to-end encryption throughout all IoT endpoints of connection and server infrastructure. Therefore, data does not get exchanged between IoT devices without proven identity.

### 2.2 *AWS IoT Rule Engine*

The AWS Rule Engine is a derived process of AWS IoT core. Rule Engine evaluates inbound data messages published IoT devices into AWS IoT core, delivers and transforms them to another IoT devices or a cloud serverless service or storage management based on IoT rules we define [4]. An IoT rule can apply to message

data from one or many IoT devices, can privilege one or many processes in parallel and route IoT rule messages to AWS endpoints such as AWS Lambda, Amazon Kinesis, S3, Amazon SageMaker. AWS IoT Rule Engine can also trigger workflows with the help of AWS Step Functions using Step Functions action. With this Step Functions action, we can customize the serverless computing using FaaS with the help of adaptable invoker to trigger the respective functions to the IoT rules.

### ***2.3 Serverless Versus (FaaS)***

Serverless computing is an execution model provided by many clouds service companies. Serverless platforms manage configuration template for server deployment behind the scenes, manage the allocation of the infrastructure and its resources [5]. Billing is based on the number of resources consumed. Serverless computing can be a form of utility computing and set which abstracts the underlying hardware for the infrastructure completely.

FaaS is a derived type of serverless computing. FaaS is quite recent thing. FaaS is very useful to handle the microservices from IoT core and narrow down to functions [4]. The core idea behind FaaS is that a small piece of code can be triggered based on an event. FaaS is one step ahead in preoccupation of servers. FaaS provides us the capability to deploy code or function in a very simple and straight method. When many microservices have to be carried in server platform, it only costs when it consumes cloud resources. Function's template can be coded in almost all the latest programming languages.

FaaS is a category of cloud computing services which is actually used in the development of applications and serverless systems. Serverless computing is a wider terminology and more concentrated toward the hardware infrastructure. But it is possible to run FaaS without using a serverless architecture.

### ***2.4 Open-Source Serverless Platform***

Open-source serverless tenets in present technology concepts have gained more demand and acceptance from developer and operational point of view. Open-source framework is mostly promoted by many companies because open-source framework adds on numerous merits to the serverless platform [9]. There are two major classifications of serverless platform, first is backend as a service (BaaS) and the second is function as a service (FaaS). BaaS features an important role in managing backend cloud servers which are running applications. By using BaaS, we can achieve simple user-friendly interface even for the complex IoT application. FaaS can run a part of code or function via an event invoker or trigger [9]. This combination of BaaS and FaaS in open-source platform permits to bring forth on-demand micro-functions and it also simplifies the solution for cloud server side IoT applications.



For the customization of serverless computing, we have chosen the open-source serverless platform known as Apache OpenWhisk which has computing power to execute functions for very large-scale response of IoT messages and events [11]. It's technically high-end serverless platform is designed using Scala language. Open-Whisk framework can execute events even through http request from private cloud. It is simple to integrate with other vendor cloud services and gives rise to optimal utilization.

2.5 Drawback in Cloud-Based Serverless Computing

Serverless computing is positioned between these two concepts:

1. The developer loses most of the control over the serverless computing infrastructure but retains control over the application. The term serverless becomes confusing for the applications where both server hardware infrastructure and server operational processes are running in two different regions provided by cloud service [6, 7].
2. Most of cloud provider organizations that are building and supporting “serverless” applications are not maintaining the hardware infrastructure or the processes. They are outsourcing this to a vendor. In many scenarios and applications, third-party tools are used for maintaining and controlling serverless computing infrastructure. As shown in Fig. 1, many background processes are carried out without the knowledge of developer and operational team.

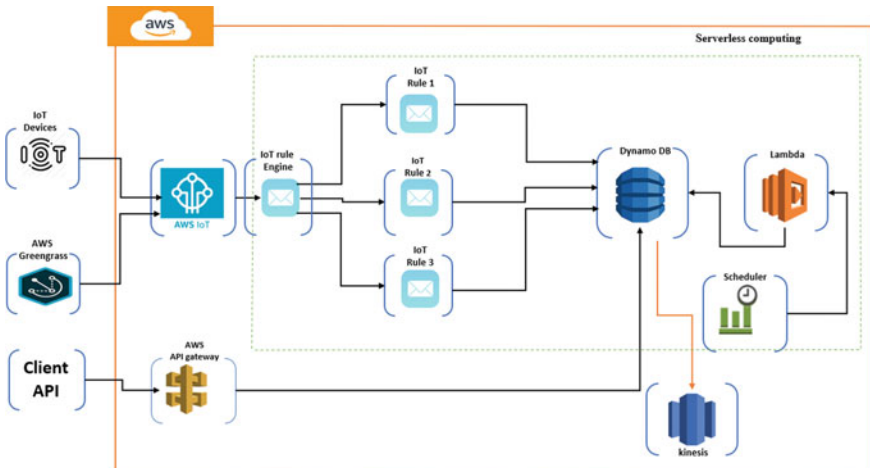


Fig. 1 Aws IoT core using lambda serverless

### 3 Lambda Serverless Service

Lambda is a serverless computing service provided by AWS cloud provider that is used to create functions and run code without managing infrastructure [2]. By lambda service, we invoke lambda function which also includes tags and metadata required for support in IoT data/message analysis—all with zero administration with respect to infrastructure [8]. Lambda has many advantages compared to the traditional manual way.

- *Application Deployment*: The speed at which developers can switch between executing code of multiple microservices is much faster than using servers such as virtual machines or containers [8].
- *Scalability*: The lambda service allows multiple concurrent executions of a function, without any operational overhead to the developer's team.
- *Low Cost*: In traditional architectures, virtual machine servers are used where we are bound to pay for the running time of an instance (virtual machines) [3], but in the lambda service we will be paying only for the time when the code is executed by a lambda service.
- *Integrations*: Lambda service can be used to subscribe various event monitoring tools like CloudWatch, SNS, Kinesis and Cloud Data Watch for data flow managements [8].
- *Stateless*: Functions used by lambda are stateless which in turn provides the simplicity in implementing and designing security systems like data analytic services that process data records individually.

### 4 Discussion

Despite the versatile capabilities of the AWS Lambda function service, there are many limitations that restrict developer design choices [3]. In this section, we describe the limits or demerits and discuss potential approaches that prevent such limitations systematically and achieve the application objective by holding majority of controls over the serverless computing.

#### 4.1 Limited Functionalities

##### 4.1.1 Event Tracing

There is a shortage of tools to trace an application event through an intricate serverless computing system to aid with troubleshooting issues and to realize the reason and where the bottlenecks exist in serverless computing system [7]. There are difficulties in handling suspicious insecure activities it is hard to identify where the suspicious

event has originated, which serverless components could be affected by the event or may be affected in future.

#### 4.1.2 More Complex Call Patterns

While using AWS lambda call patterns, each function code is predetermined with process time on the assigned containers in the back-end infrastructure. When the execution time increases more than the predetermined process time, the developers are not able to run the particular functions code for the application. Most of the lambda functions have default time of 3000 s. As a result, maximum time is spent on orchestrating the lambda serverless platform by the developers. Distributing the function code for complex application becomes a challenge with lambda call patterns [6].

When tasks involving large amounts of data are processed using lambda call pattern, it will mostly have the habit of overshooting the default process run time. Now the developers have to put significant effort in orchestrating the rewritten function code and assigning the new process time for each function possibly using a different infrastructure for function executions. This leads to complex call patterns in serverless computing platform [6].

When the endpoint inputs information given to lambda function, AWS lambda proceeds in creating a temporary virtual container, deploys all dependencies data or messages and runs the required code. After the request is executed and completed, the container is destroyed from the infrastructure platform and is not suitable for real-time application.

However, starting up or creating this temporary container can take between 100 ms to 120 s and this delay in execution is known as a cold start.

#### 4.1.3 No Control Over Infrastructure

Many IoT applications in heterogeneous mashup architecture will need different platform operating systems simultaneously [6] to execute the function codes. Under these conditions, the application has to be migrated or hosted in different infrastructures which are redesigned for respective operating systems.

Some tools are provided by AWS to preserve consistency in monitoring the code/function execution process but it is not assured that the same infrastructure configuration will be received for every execution in different operational platforms. Finally, lambda service does not provide the developers control over instances that serve in serverless platform.

#### 4.1.4 Disk Space

AWS limits the size of disk space under the “/tmp” directory related to a lambda function which in turn restricts using lambda for workloads like big data analysis and video encoding. Even documentation is not provided about lambda disk’s encryption. Due to this, the developer finds difficulty in mounting a disk with different encryption technology.

## 5 Serverless Computing Using Apache OpenWhisk

In cloud platform, we have many serverless services from giant cloud providers like AWS, Azure and Google but Apache OpenWhisk is open-source software solution provided by Apache. It has contributions from multiple companies in open-source software development which results in user-compatible serverless infrastructure. IBM, RedHat and Adobe cloud platform APIs are multiple public providers [10] which gives free hand for the developers to customize the serverless infrastructure in different operating platforms to achieve the objective of the client requirements. It also provides self-host deployment model [9, 11]. Figure 2 shows concept of DevOps style implementation by customizing serverless cloud platform using Apache OpenWhisk and FaaS deployment. This prototype model helps in minimizing the drawbacks experienced by the developers and operational team working on serverless platform.

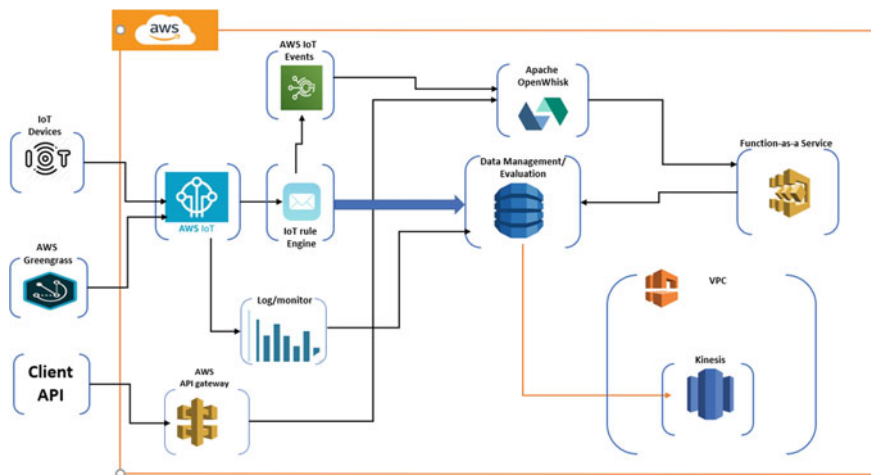
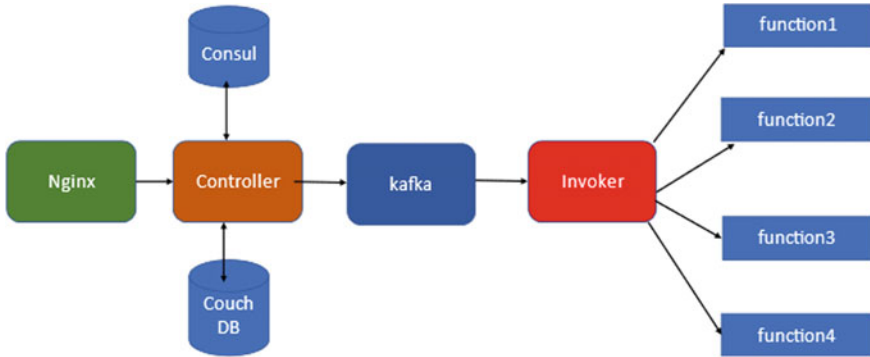


Fig. 2 Implementing DevOps Style in Serverless using FaaS and Apache OpenWhisk



**Fig. 3** Building components of OpenWhisk

By using AWS API gateway with http client, we can route the action to be carried out by OpenWhisk and provide custom domains security with API keys. OpenWhisk is multiplatform portable serverless engine between different commercial cloud vendors [9]. As shown in Fig. 3, the OpenWhisk serverless infrastructure keeps waiting for any important event to turn out. The function gets triggered or invoked only when event turns out.

### 5.1 *Apache OpenWhisk Operation*

For the developers and the end-users, the OpenWhisk operation is very uncomplicated and plain sailing but several internal components installation are carried out automated before the user can view the OpenWhisk platform [9, 11]. The internal components of OpenWhisk are shown in Fig. 3. Asynchronous method of process is adopted in OpenWhisk, which gives a friendly hand for clients to start requesting many events without waiting for the acknowledgements. Client receives in return an activation ID from the invokers. Internal databases are used to store the activation ID and its parameters. When process or event is completed, the result of execution, logs and other information can be fetched from the internal database.

- **Nginx** This internal component of OpenWhisk falls under the group of open-source web reverse proxy server. Using https endpoints clients can hit the request to the OpenWhisk framework [11]. The Nginx layer can be easily scalable up/down according to the requirement.
- **Controller** It performs the authorization and authentication on every entity and routes the action. It takes the role of a sentinel of the OpenWhisk platform. It plays an important role in managing the trigger.
- **Apache Kafka** It facilitates in building distributed, scalable, and publish-subscribe message/event infrastructure [11]. It provides a backbone in handling billions of event ingestions.

- **Invoker:** It takes the work of setting up the containers and pushes the respective functions for execution. When the execution is completed, the invoker pushes the log and information data into the couch database for upcoming event usage and also for future referral.

## 5.2 Drawbacks in the Proposed System

The DevOps team working on components of OpenWhisk needs to be highly skilful in python and Nodejs because it is a rich programming model. It also requires grip understanding about the role of containers in function as a service. A function is used to call another function which in turn gives rise to toil in testing and monitoring. The invoker cannot handle the change in functions dynamically and redeploy.

## 6 Future Work

Even after the work accomplished till now, we consider that there is still a set of opportunity open for many research challenges that are worthy to be explored in the coming years. Serverless computing goes even beyond cloud computing. Therefore, it is one step ahead which again opens opportunity to work with many open sources distributed serverless platforms such as OpenFaaS, Knative, Firecracker and Oracle FN. We can work on to replace the dynamo DB used for Data management and evaluation with Couch DB so that the monitoring becomes simple using Apache OpenWhisk [9]. By designing even-driven application and embedding with invoker by using the benefits of both trigger and queues, the difficulty in testing/monitoring and change function redeployment can be made dynamic, in turn achieve deployment of functions with lower cost with custom rules.

## 7 Conclusions

To ease the development of serverless services in the cloud, this paper describes serverless DevOps design patterns that can be used to build serverless applications for IoT cloud platform services. We have highlighted the key advantages of lambda services and its major setbacks in functionalities and difficulties faced by developer in serverless applications. We also conceptually explained that a large-scale of data processing from IoT devices in real time is difficult by the proposed design rules, cloud provider patterns in serverless and FaaS system by cloud providers. In addition, we described the limits of lambda functions and provided ways to overcome them by using open-source Apache OpenWhisk [11]. We envision that the proposed serverless design patterns will revolutionize IoT application deployment, monitoring,

increase the data message throughput and security systems in the cloud serverless platform with private facilities even when the infrastructure is running on public cloud platform. Despite the growth of many open-source serverless platforms in the technology market, Apache OpenWhisk will continue to gain popularity among the developers in FaaS and serverless platform [11]. It is worth noting that these platforms can be easily compatible with hybrid operating models because they designed by IBM, RedHat, and Adobe cloud platform with different platform open-source software code. Therefore, these attributes will lead for more and more research in this domain of serverless.

## References

1. Akkus IE, Chen R, Rimac I, Stein M, Satzke K, Beck A, Aditya P, Hilt V (2018) SAND: towards high performance serverless computing. In: Proceedings of the USENIX Annual Technical Conference (USENIX ATC)
2. Khvoynitskaya S (2020) The history and future of the Internet of Things. Available online: <https://www.itransition.com/>. <https://www.itransition.com/blog/iot-history>. Accessed 25 Mar 2020
3. Monther AA, Tawalbeh L (2020) Security techniques for intelligent spam sensing and anomaly detection in online social platforms. *Int J Electr Comput Eng* 10:2088–8708
4. <https://docs.aws.amazon.com/iot/>. Accessed 21–2–2021
5. Meng Y, Zhang W, Zhu H, Shen XS (2018) Securing consumer IoT in the smart home: architecture, challenges, and countermeasures. *IEEE Wirel Commun* 25:53–59
6. Siby S, Maiti RR, Tippenhauer NO (2017) Iotscanner: detecting privacy threats in IoT neighborhoods. In: Proceedings of the 3rd ACM international workshop on IoT privacy, trust, and security, Abu Dhabi United Arab Emirates, 2 April 2017, pp 23–30
7. Hassan WH (2019) Current research on Internet of Things (IoT) security: a survey. *Comput Netw* 148:283–294
8. Amazon lambda. <https://aws.amazon.com/lambda/>. Accessed 2021–2–7
9. <https://projects.apache.org/committee.html?openwhisk>. CrossRef
10. Cloud Controls Matrix. <https://cloudsecurityalliance.org/group/cloudcontrols-matrix/Overview>, 2018. Accessed 03–2–2021
11. Apache Software Foundation, Apache OpenWhisk. <https://openwhisk.apache.org>

# A Cloud Interfaced Social Ski Driver Optimized Routing for Future Electric Motor Vehicles



Shivanand C. Hiremath and Jayashree D. Mallapur

**Abstract** India is revamping towards building smart cities along with collaboration of information and communication technologies as a future mode of transportation and with a vision of adapting electric motor vehicles (EMVs) more than 90% to make electric vehicle nation by 2030. With scarcity and increased floating prices of combustion fuels, it has procured enormous intrigue among various researchers as the EMVs are termed as an upcoming incredible mode of transportation. In addition, EMVs protect environment by satisfying go-green initiative and it is vital to protect conventional fuels and use renewable sources. However, the EMVs have shorter driving ranges, which are limited by inadequate charge storage in batteries, EMVs are economical compared to using gasoline but due to irregularly distributed charging stations causing lack of interest and hindrance among people in adoption of EMVs. This paper presents unique strategy to assist EMVs users through optimum routing directions to charging stations with ease of access by using cloud hosted on VANET. The goal is to locate the shortest and with least traffic routes for charging EMVs through the cloud-based vehicular ad hoc network (VANET) model. Here, optimum routes for acquiring the nearest charging station find out using proposed social ski driver (SSD) algorithm and a comparative analysis is done with particle swarm optimization (PSO) algorithm by considering traffic density, battery power and distance parameters. The experimental result of proposed SSD algorithm outperforms the PSO.

**Keywords** Electric motor vehicle (EMVs) · SSD · PSO · VANET · CS

---

S. C. Hiremath (✉)

R. N. Shetty Polytechnic College, Belagavi, Karnataka, India

J. D. Mallapur

Basaveshwar Engineering College, Bagalkot, Karnataka, India



# 1 Introduction

In subsequent days, modernization of India is having a vision of smart city development and eco-friendly go-green initiations leading to growing heed towards usage of EMVs instead of inflammable fuel-based vehicles due to pragmatic impact on the environment by reducing CO<sub>2</sub> emissions, as well as to overcome the routine price fluctuations of inflammable fuels. However, India's increasing population rate demanding and exceeding the import dependency of heavy crude oils from other countries for consumption every year and its variations in costs become a major setback for a common man, that is, leading to socio-economic crises. Instabilities of fuel costs and their sustainable deployment require new technologies such as refining and enhancing battery technologies and making the batteries to adopt to turbo recharge facilities and also needs a new cloud-based platform for dynamic provisioning of charging stations (CS) with time-efficient recharging of the EMVs during long drives. In spite of many researches are proposed on EMVs charge scheduling, and various methods of interfacing with charging resources and traditional techniques are provided with road traffic assessment on EMV planning paths. Major unaddressed challenging issue is to minimize the impacts of charging vehicles to grid load at charging stations particularly at busiest streets. In India, market of electric motors is quite discouraging due to challenging issues such as charging process of vehicles such as time consumption to charge, lack of ease of availability of charging station, increase in the peak load period, decrease in service quality, degradation of the voltage profile, overload of circuits, and increase in energy losses. Government of India is encouraging to setup the public charging stations (PCS), and EMVs, major constriction, are their batteries that often consume very large time span depending on their capacity and to run for large mileages. Therefore, deploying charging stations that can properly coordinate the charging process of EMVs Hence, drawbacks and issues confronted by traditional techniques are addressed to locate the nearest charging station by considering available grid power at charging station, vehicle battery status, distance to reach and road traffic conditions to encourage ease of access to the people to adopt EMVs. A challenging issue of dynamic provisioning of charging station, inspired by futuristic automotive market consisting of electric motor vehicles and its refuelling charging station in metro cities as well as maximizing the utilization of the charging station resources while satisfying the charging demands of the EMVs by finding the fitness value considering battery status ( $B_{soc}$ ), traffic density ( $T_{den}$ ) and distance parameters to assist optimal shortest route with less traffic from the current vehicle position by maximizing battery efficiency. In this paper, we propose a unique and evolutionary social ski driver (SSD) optimization algorithm to make dynamic real-time routing decisions by finding the best optimal fitness value to assist users for charging on the go.

## 2 Related Works

The coordination of EMVs and its charging within a station raises challenging issues, and many charging models with scheduling approaches have been come up with a panoramic review of charging. Decentralized recharge allows deciding the period to start charging their EMVs, whereas centralized recharging takes decisions in the charging station by using requested information provided by the EMV owners. Although the centralized cloud scenario is more secure than the decentralized, as the overall system can be controlled in such a way to prevent overloading of the power grid and also minimizes power losses. Moschoyiannis et al. [1] delineate about dynamic wireless charging of electric vehicles on the move using spring coils without recharging infrastructure, heavy vehicles like buses or trucks act as mobile energy disseminators this dynamic idea can pose health hazards and impractical to carry loads of batteries to provide recharging facilities. Miguel et al. [2] discuss charging parking lots to satisfy user demands with available power but the major dynamic problem is to make user to know the locations of refuelling stations needs to be addressed. Afida et al. [3] outline about types of battery chargers and its associated equipment's, a detailed study on the commercial and prototype electric vehicles for the future in terms of electric range, battery size, charger power and charging time required. Behnam et al. [4] propose an optimal electric vehicle charging scheduling (EVCS) in the distribution grids, and this paper only deals with optimization of energy loss and electricity cost through centralized operator. Michalis et al. [5] discuss scheduling problem in a charging station to address the coordination of the charging process for fleet of EVs but this work needs to handle realistic constraints such as variable charging rates as per user demand with dynamic location to improve the performance of ACO. Huang et al. [6] delineate about plug-in hybrid electric vehicles (PHEV) that charges on demand vary stochastically with time and have only considered scheduling of the time with a constant charging rate Siam et al. [7] demonstrate the FPGA-based parking management system and use deep neural convolution network algorithms, and the images are processed to identify unoccupied spaces, neither it provides charging with parking dynamicity nor economical solution to use in real-world traffic conditions. Shyam et al. [8] propose Internet of things (IoT) with scalable and cost-effective solution to indicate unoccupied parking spaces. It uses Raspberry Pi and infrared photoelectric reactive sensors to identify the presence of object by sensing distance, and hence, this system checks the space and updates to cloud server but does not informs the shortest time and distance path guide to parking.

## 3 System Overview

The idea of proposed architecture is derived from the India's subsequent vision of smart city development plans and go-green initiatives.

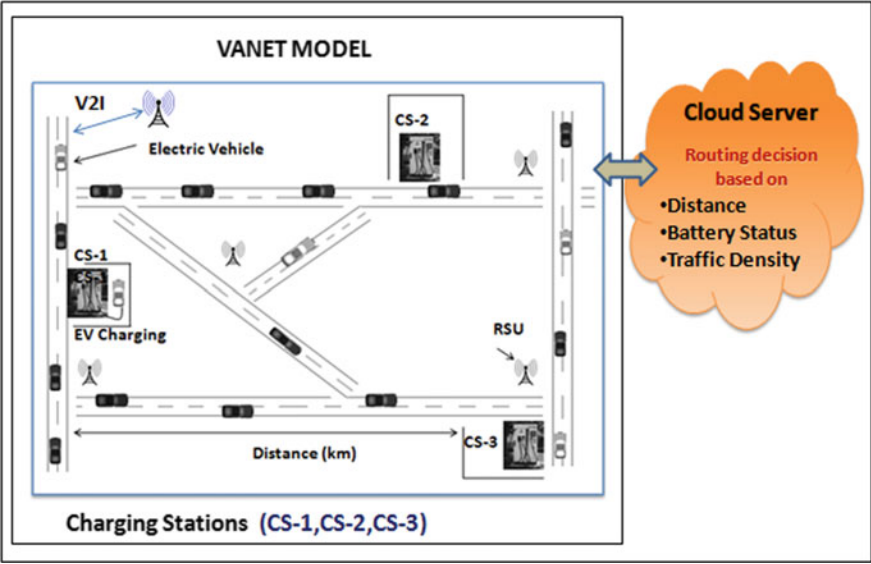


Fig. 1 System overview

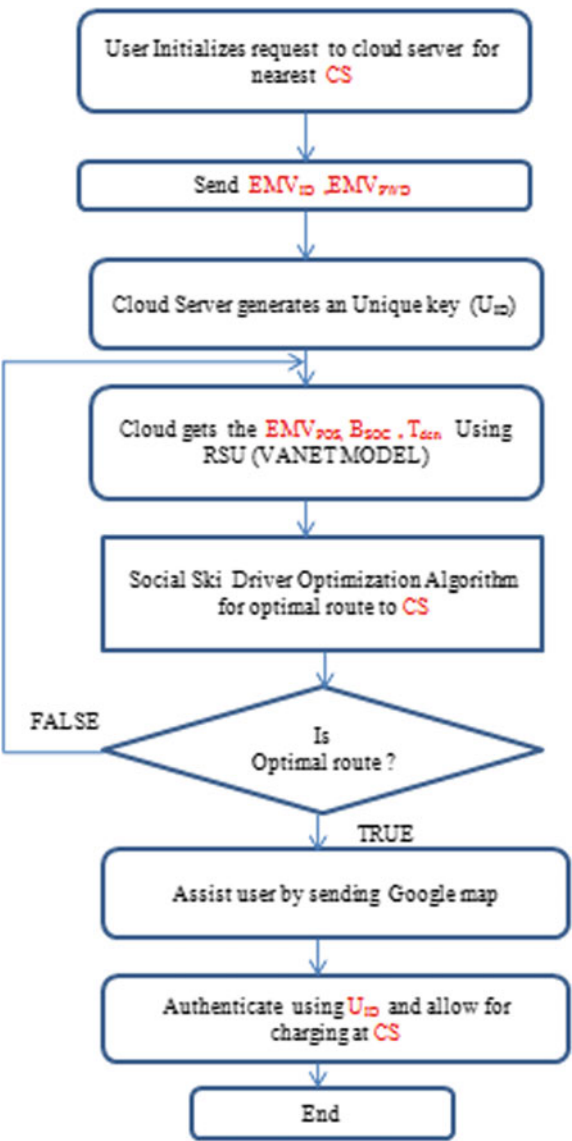
VANET model in Fig. 1 provides different communication scenarios such as vehicle to infrastructure (V2I), RSU to central cloud server. When the request is made from the user for CS, cloud server gathers traffic density ( $T_{den}$ ) information from VANET model and EMVs sends its data such as current position ( $EMV_{pos}$ ), battery status ( $B_{soc}$ ). The proposed model considers the available grid capacity at CS and then finds the optimal solution by running SSD routing algorithm in cloud server through multiple iterations to get the best mean of the fitness value to select the optimal route considering min distance, less traffic density and minimum battery power needed to reach the nearest charging station (CS).

3.1 System Flow

Figure 2 explains the overall system operations for computing the optimal solution for routing electric motor vehicles (EMVs).

EMVs have their own driving ranges based on their battery capacity EMV owner has to decide for long or short journey based on current status of the battery, and we never know when we come across to go for long drive beyond the driving ranges of EMVs to overcome such a panic situation we have proposed an architecture to support the charging of EMV; whenever battery power reaches threshold value after a certain driving range, a request message consisting of  $EMV_{pos}$  and  $EMV_{PWD}$  is sent to cloud server to find the nearest charging station at the earliest.

**Fig. 2** Flowchart for system operation



Due to fixed driving ranges of EMVs, each metre of movement and idling of vehicle also consumes the power under traffic congestions, hence computing of the shortest route along with minimum traffic density becomes one of the vital issue. We have come along with design of computing fitness values considering the current battery status ( $B_{soc}$ ), minimum traffic density ( $T_{den}$ ) and minimum distance to reach nearest charging station from the requested position of electric motor vehicle. Traffic density is evaluated from the data taken through Google Maps and local VANET

models consisting of roadside units that synchronize with central cloud server every fixed time. Battery status of charge is computed using Coloumb's counting method that measures status of charging, and discharging current of the battery by integrating the values over time is given by

$$B_{\text{SOC}}(t) = B_{\text{SOC}}(t - 1) + \frac{i(t)}{A} \alpha t \quad (1)$$

where  $B_{\text{SOC}}(t)$  represents battery state of charge at time instant ' $t$ ',  $B_{\text{SOC}}(t - 1)$  denotes initial state of charge at time instant ' $t - 1$ ' and  $i(t)$  battery current at time instant ' $t$ ',  $A$  is the capacity of the battery in ampere hours (Ah) and  $\alpha t$  denotes step time.

Electric motor vehicles are one of the futuristic automotive markets as an alternative solution to overcome the scarcity and increased floating price of gasoline day by day, and also, there exists a lot of challenging issues such as lack of encouragement and motivation towards the usage of EMVs among people, agitation about battery-operated vehicles, difficulty in availability of charging station due to oblivious reasons; hence, these issues can be handled dynamically using application to assist users to have ease of access to charging stations during short, medium and long journeys while on the go in metro cities and urban areas this can be achieved using social ski driver algorithm.

### 3.2 Social Ski Driver Algorithm

Social ski driver algorithm is inspired by numerous metamorphic techniques used in finding optimal values for support vector machines [9] and also used in feature selection [10]. Here the aim of SSD is to search the space for getting near to the best of the optimal solutions from the past data records through multiple iterations or simulation rounds.

The SSD computes the best optimal solution as follows:

*Position of the EMVs:* The current position of EMVs is computed as objective function at that geographical space.

*Previous best position obtained:* For each EMVs, the computed fitness value is compared with previous best data records and stored as foremost position of EMVs.

*Mean global optimal solution:* Finally to find the global optimal solution by taking the mean of all fitness values are calculated.

---

***Social Ski Driver (SSD) Algorithm:***


---

- ***Input:*** Number of EMVs  $N$
- ***Output:*** Best optimal solution  $N^*$
- ***begin***
- ***Initialization of EMVs Current Position and Velocity***
- ***while*** Stopping criterion not met ***do***
- ***for all*** EMVs ***do***,
- ***Get the battery status of EMV using Eqn(1)***
- ***Estimating values of fitness .***
- ***Sorting out EMVs based on fitness values.***
- ***Evaluate past best solution & Mean global solution.***
- ***Compute new optimal solution .***
- ***Alter Velocities of EMVs.***
- ***End for***
- ***End while***
- ***Return optimal solution  $N^*$***

## 4 Results and Discussions

The analysis of existing population-based particle swarm optimization (PSO) is an evaluation technique where each particle decides its movement through search space by combining its present and best fitness functions to compute its final fitness values at its current location; hence, it is used as a routing protocol in various researches so we have used PSO to compare with our proposed social ski driver (SSD). Considering 5 EMVs requesting for charging stations for simulation, we have considered a general traffic of 100 vehicles randomly distributed on lanes. The performance metrics are evaluated based on four parameters traffic density, delay, power and fitness values, and a number of simulation rounds or iterations are repeated to acquire the best optimal solution.

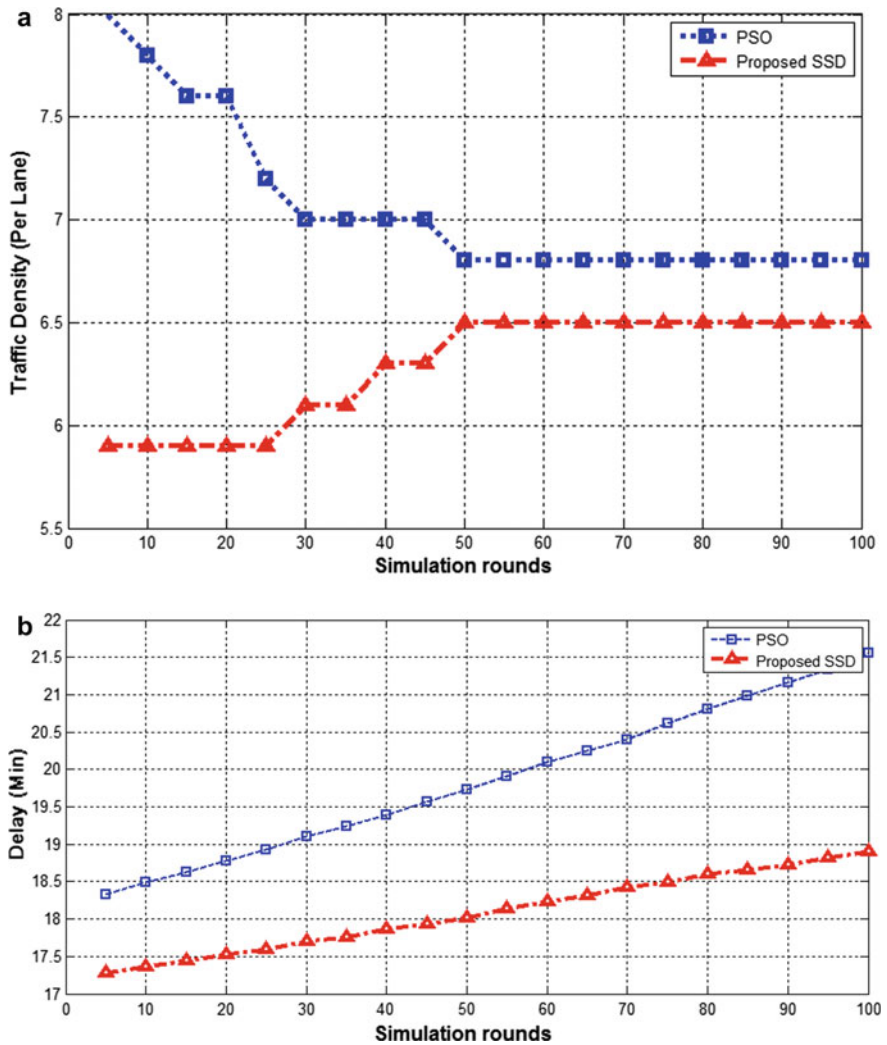
### 4.1 Experimental Setup

Simulation is done using Windows 10 OS using Intel core i5 processor and 4 GB RAM and executed in Python. A metaheuristic optimizing technique particle swarm optimization (PSO) is used for comparative analysis with proposed SSD algorithm.

The analysis of methods with traffic density is shown in Fig. 3a. For 20 simulation rounds, the traffic densities measured by PSO are 7.6 per lanes and proposed SSD are 5.8 per lanes similarly for 100 simulation rounds, traffic density per lane are

measured by PSO are 6.8 and proposed SSD are 6.5; hence, the proposed method selected the route with minimum traffic per lane.

Figure 3b shows for 20 simulation rounds, the delay in computing the optimum routes by PSO and proposed SSD are 18.727 min and 17.418 min, respectively; similarly, for 100 simulation rounds, the delay measured by PSO and proposed SSD are 21.898 min and 18.898 min, respectively. Hence, the proposed method is shown an improved performance with minimum delay helping EMVs to reach charging station (CS) at earliest.



**Fig. 3** a Analysis of traffic density, b analysis of delay (per Lane), c analysis of battery power (W), d fitness values after multiple iterations

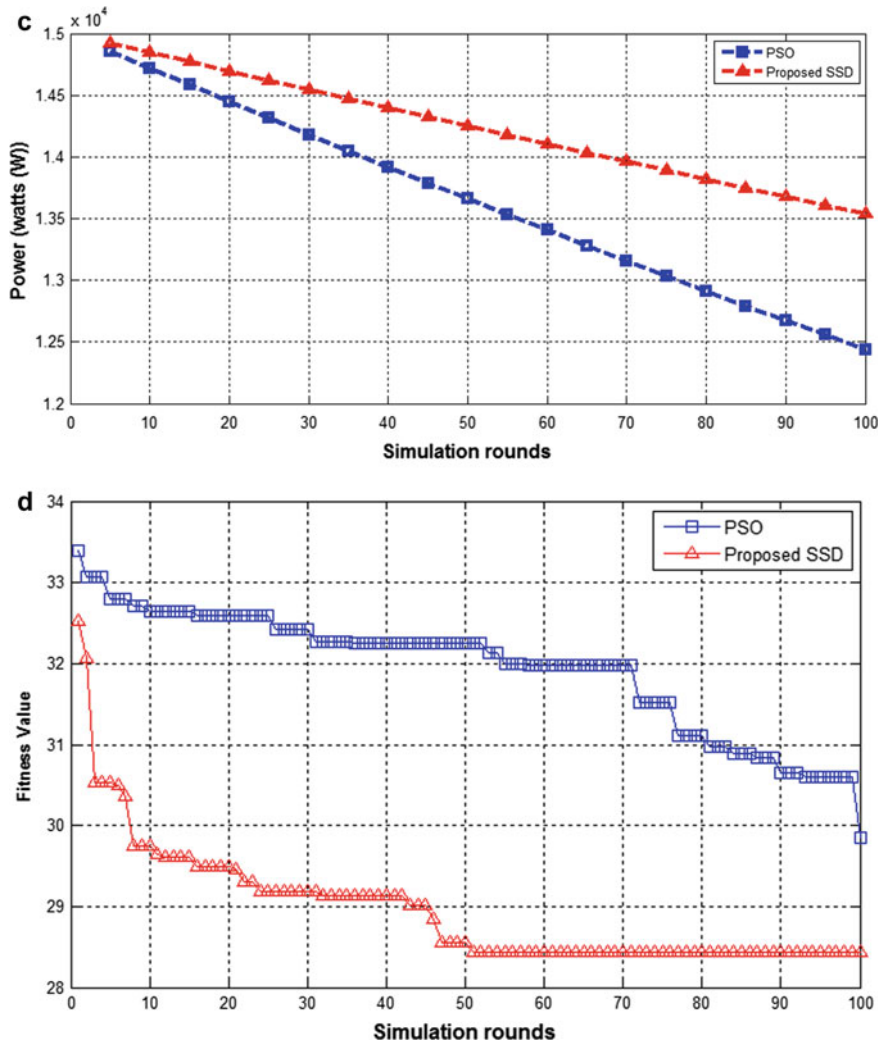


Fig. 3 (continued)

Figure 3c denotes the analysis with battery power consumed after reaching charging station (CS) using both PSO and Proposed SSD optimal routes. For 20 simulation rounds, the available power measured by PSO and proposed SSD are 14,760.84 W and 14,370.63 W, respectively; hence, 390.21 W of battery power has been saved, and similarly for 100 simulation rounds, the power measured by PSO and proposed SSD is 13,540.98 W and 12,584.94 W, respectively, with 956.04 W battery power savings. Hence, minimum usage of battery power to reach has been improved performance of battery efficiency.



The fitness values are evaluated considering multi-parameters battery power, traffic density and distance as shown in Fig. 3d, for 20th iteration, the fitness measured by PSO and proposed SSD are 32.589 and 27.118, respectively. Likewise, for 100th iteration, the fitness measured by PSO and proposed SSD are 28.440 and 24.800. Hence, the computed fitness values for various simulation rounds are considering traffic density, battery power status and delay outperforms the PSO optimization technique.

## 5 Conclusion

This paper proposed an unique concept that utilizes the centralized cloud-hosted VANET model to route the EMVs to reach the nearest charging stations in short time with higher efficacy, and the proposed SSD takes less time to find destination from current position and offered a better quality of solution; hence, SSD routing has the advantages over PSO routing technique and can be adapted to select shortest optimum route for charging EMVs, by increasing battery efficiency. The proposed SSD optimization enhanced the performance by using minimal traffic density of 6.5, minimal delay of 10.973 min, and battery power efficiency and minimal fitness of 24.800, respectively. In future, furthermore comparative analysis among metaheuristic optimization techniques needs to be done to check its efficiency for predicting the nearest charging station.

## References

1. Moschoyiannis S, Maglaras A, Argyriou A, Janicke H (2018) Route optimization of electric vehicles based on dynamic wire-less charging. *IEEE Access* 6:42551–42565
2. Gonzalez MA (2018) Electric vehicle charging scheduling by an enhanced artificial bee colony algorithm. *Energies* 2018. <https://doi.org/10.3390/en11102752>
3. Faizal WM, Mahmood W (2014) Review on electric vehicle, battery charger, charging station and standards. *J Appl Sci Eng Technol* 7(2):364–373. ISSN: 2040-7459
4. Khaki B, Chung Y-W (2019) Hierarchical distributed EV charging scheduling in distribution grids. In: *IEEE PES General Meeting*, [arXiv:1812.02847v2](https://arxiv.org/abs/1812.02847v2) [math.OC] 27 Mar 2019
5. Mavrovouniotis M, Ellinas G, Poly-carpou M (2019) Electric vehicle charging scheduling using ant colony system. *Conference Paper May 2019*. 10.1109. <https://www.researchgate.net/publication/333055684>.
6. Huang J, Vijay Gupta, Huang Y-F (2012) Scheduling algorithms for PHEV charging in shared parking lots. In: *2012 American Control Conference (ACC) IEEE Conference* 27–29 June 2012. <https://doi.org/10.1109/ACC.2012.6314939>
7. Siam M, Valipour S (2016) Parking Stall vacancy indicator System based convolutional neural Networks. In: *Internet of Things (WF-IoT) 2016, IEEE 3rd Worlds Forum*, pp 655–660
8. Shyam JR, Nrithya T (2017) Cloud Connected Smart Park International conference on I-SMAC (IoT in Social Mobile, Analytics and Cloud)
9. Tharwat A, Gabel T (2019) Parameters optimization of support vector machines for imbalanced data using social ski driver algorithm. Springer-Verlag London Ltd, part of Springer Nature. <https://doi.org/10.1007/s00521-019-04159-z>

10. Chatterjee B, Bhattacharyya T, Ghosh KK, Singh K, Geem ZW, Ram Sarkar. Late acceptance hill climbing based social SkiDriver algorithm for feature selection. <https://doi.org/10.1109/ACCESS.2020.2988157>. IEEE Access

# Cloud-Enabled Education-as-a-Service (EaaS)—A Review



Shubhpreet Kaur, Tarandeep Kaur, and Ashok Sharma

**Abstract** Education plays a significant role in sustaining the economic growth of every country. Learners are becoming more technology-based and more advanced than ever before; learning and teaching are changing day by day. Cloud computing is transforming the ways of individual learning and interactive learning. The effective teaching standards provide quality and interactive training such as e-learning to students and smart training to teach students. Information technology (IT) plays an important role in the education field. Nowadays, e-learning and smart learning have become very popular trends in education technology. E-Learning is the new model that provides digital content and an education-oriented environment for teachers and students using Internet technology. Smart learning is a model of learning that has been adopted for new generations of digital natives. This paper discussed the concept of cloud computing for e-learning and smart learning. The paper also describes delivering EaaS through the use of cloud computing and also discusses the benefits of EaaS.

**Keywords** Cloud computing · Education-as-a-Service · Information technology · e-learning · Smart learning · Software-as-a-Service · Infrastructure-as-a-Service · Platform-as-a-Service · Massive open online courses

## 1 Introduction

The education system is in a constant process of change; institutions have to adapt to the needs, aspirations, and essentials of the learners. Thus, just as information technology and E-learning systems are seen as important factors to complement the activities of institutions and organizations, these organizations and universities are

---

S. Kaur (✉) · T. Kaur · A. Sharma  
Lovely Professional University, Punjab, India

T. Kaur  
e-mail: [tarandeep.24836@lpu.co.in](mailto:tarandeep.24836@lpu.co.in)

A. Sharma  
e-mail: [ashok.23877@lpu.co.in](mailto:ashok.23877@lpu.co.in)

increasingly investing in online systems and devices. The use and need for IT support including servers, storage, and software is multiplying in universities, colleges, and schools. IT support is concerned with providing computing technology and infrastructure (such as hardware, software, networking, and internetworking). Many education systems have IT departments to cope up with and manage the technological advancements of computers, networks, and other businesses. The dissemination of education using IT can aid in promoting knowledge dissemination opportunities for teachers and students along with the complement and implementation of the latest technological innovations [1].

Presently, delivering education online is the need of the hour especially when the global norms for social distancing and virtual interactions are rising. IT has made the education process more efficient and beneficial thereby increasing the well-being of the students. The use of technological training methods such as replacing books with electronic books and devices (tablets and laptops) has simplified the education process. These platforms can be a real option for people who are out of school, or who have difficulty connecting with their teachers in the classroom [2]. These platforms give students the opportunity to review courses at every moment with simpler and more concise explanations, and this strengthens the process of education and in schools for most students brings better results [3].

The current pandemic crisis has emerged as a boon for the virtual learning and e-learning platforms. Virtual learning and e-education platforms enable the students to study from their home environment specifically in such crises.

### ***1.1 E-Learning***

E-learning is an education concept that provides digital content and an education-oriented environment for teachers and students using Internet technology. Internet, audio/videotape, CD-ROMs, and many more are used to download and load contents. It supports learning and teaching through the use of Web technology and computer systems. Advances in the Internet and multimedia technology enable e-learning. E-learning bridges the gap between a learner and an educator in two different geographic areas. It comprises all forms of electronically supported learning and teaching (like audio, video, multimedia communication, site, and online classrooms) [4].

### ***1.2 Smart E-Learning***

Smart e-learning is a new model for e-learning that has teachers engages the learners to demonstrate knowledge and achieve satisfactory results in their learning process. The platform intervenes between the teacher and the learner. The teacher-learner relationship is formally designed to define the needs of the learners and to value the efforts of the teachers to meet these requirements [5].

Subsequently, the demand for delivering e-learning has been upgraded with the aid of novel electronic and technological gadgets, devices, and platforms leading to the growth of new learning models, that is, smart learning. Smart learning is “a model of learning that has been adopted for novel generations of digital nations [6].” Compared to traditional models, smart learning is an interactive, collaborative, and visionary model designed to enhance learner commitment and enable educators to adapt to learner’s abilities, interests, and educational preferences. Smart learning supports the professionalization of the education process, supporting teachers to better prepare and improve their speeches and respond flexibly to the requirements of learners and situations in the classroom, leading to increased skills and better teaching. In the learner-centric context, the teacher is a guide, an effective two-way learning partner [7].

Recently, cloud computing has emerged as an innovative technology for offering e-learning and smart learning at a higher level [8].

## 2 Cloud Computing

Cloud computing offers an important proposition that differs from the traditional enterprise information technology (IT) environment. It enables global access to cloud computing services and information with minimal investment through a flexible convenient computing environment that supports demand scalability. Cloud computing can also offer pre-built explanations and facilities, with the help of the skills needed to operate and maintain them, possibly minimizing risk and retaining a less skilled workforce. Cloud computing is the delivery of computing services—including servers, storage of data, networking, software, analytics, and intelligence—to the online to provide rapid innovation, flexible resources, and economies of scale. You usually only pay for the cloud services you use, which helps reduce your operating costs, run your basic operations more efficiently, and scale your business as needed [9]. It is important to understand the different types of cloud computing services. Below are the various services provided by cloud computing [10] (Figs. 1 and 2).

### 2.1 *Smart and Virtual Learning Through Cloud Computing*

Cloud-based e-learning is virtual learning that takes place on the cloud—an essential space that is not tied to anybody’s computer [11]. An elastic and Internet-based data storage hosting service allows e-learning providers to host their proprietary software over a centralized network through cloud computing. The cloud permits learners to access e-learning content from everywhere with a network connection. There are several advantages to using cloud computing for e-learning [12] (Fig. 3).

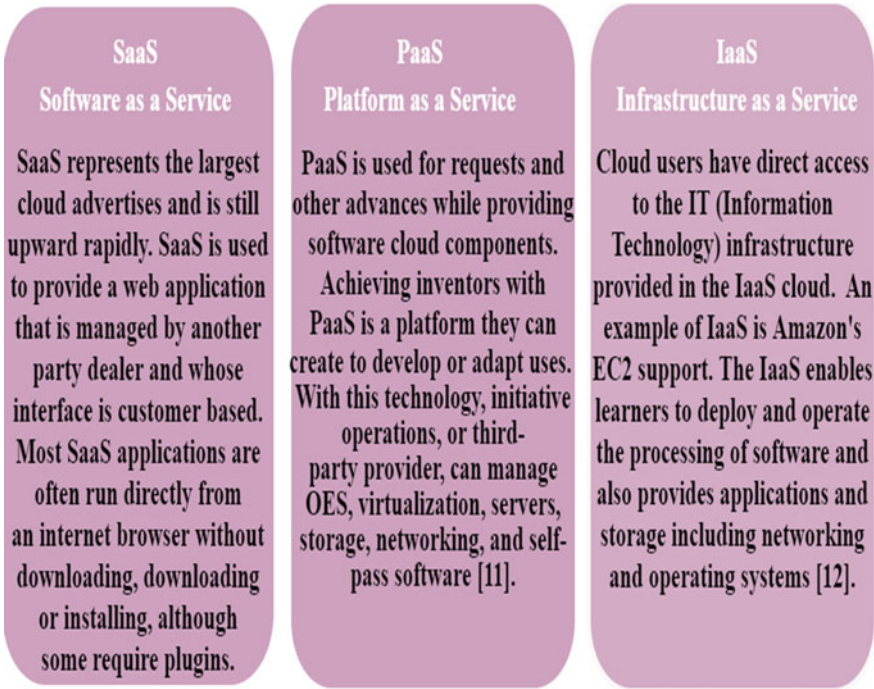


Fig. 1 Services of cloud computing

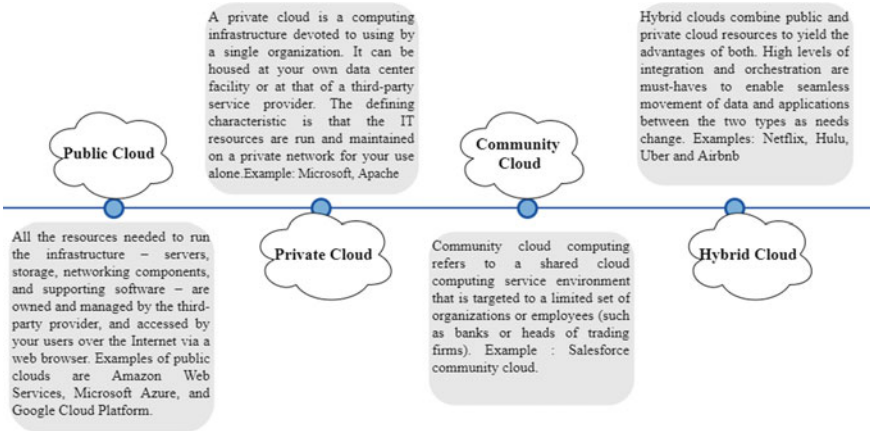


Fig. 2 Summarization of cloud deployment models

Learning Anywhere,Anytime	Learners can access learning materials anywhere and anytime. It enables independent education and makes it easy for distance students to fully join in the course.
Provides different learning styles	Cloud platforms facilitate learners and teachers to the use of videos, interactive competitions, learner-teacher discussion environments, online tests assignments and more.
Self-paced	Learners can learn at their individual speed. It improves fulfilment and reduces the level of stress.
Reusable Learning Materials	Learners can prepare assignments and exams from previous lectures given by teachers in online classes. Learners can reusable the learning materials one or more times.
Feedback	Teachers can get immediate feedback from learners as well as send responses for development to their students.

Fig. 3 Benefits of cloud computing for e-learning

Apart from offering SaaS, PaaS, and IaaS, cloud computing has elevated to delivering more services such as database-as-a-service, healthcare-as-a-service, and education-as-a-service.

### 3 Education-as-a-Service (EaaS)

The EaaS model provides learners with another or enhancement to more expensive four-year degree programs and delivers customized learning opportunities for learners. The EaaS model facilitates educational organizations and businesses to utilize courses that align with their offering without having to get excessive learning materials that are not being executed in its program. EaaS is personalized because each learner has access to the most important learning areas for them and can pursue them at their own pace. It is cheaper because it is priced according to individual courses and content levels, not a full program or degree course. It is scalable because, like software by the cloud, organizations can quickly access business groups, geographies, and job classifications, enabling greater capacity [13].

#### 3.1 EaaS Learning

EaaS learning is defined as providing learning and development to learners/teachers through various electronic media such as the Internet, audio, video. EaaS learning is

the latest and most popular form of distance learning today. Within the last decade, it has had a major impact on post-secondary education, and the trend is only growing. EaaS learning is education that is done through the Internet. This is often called “e-learning” in other words. Some benefits of EaaS learning are [14]:

- **Accessibility:** With no geographical boundaries involved, accessibility is one of the key considerations. For some learners, moving to another country for study purposes is not the only option. EaaS learning removes all restrictions that allow learners around the world to complete the course of learning in which they are interested.
- **Resource scalability:** With EaaS learning, a teacher’s reach is almost infinite. By saving money and time, other resources can also be preserved. Scalability enhances the effectiveness of learning and teaching processes, with the support of both parties.
- **Self-paced learning:** Online courses allow learners to enjoy a self-paced learning program. They can study according to their schedule, regardless of their busy schedule; it is easy to incorporate learning into their busy lives. Full-time employees who do not have enough time to attend physical classes are no longer left behind. After work, in the evenings, on weekends, or when they have free time, it is possible to learn and complete courses after this course through this digital mode [15].
- **Effectiveness:** EaaS learning has a positive effect on an organization’s profitability. It makes it easier to understand content. It results in improved scores on certificates, tests, or other types of assessments. The highest number of students achieving a “pass” or expert “level” to enhance the ability to learn and apply new developments or awareness in the workplace which help maintain information for a longer period of time.
- **Promotes active learning:** An EaaS learning environment supports active learning. The online educational videos are designed to help students stay connected to the topic of study, solve problems in a variety of ways, and form a forum to discuss problems, new ideas, and more.

### ***3.2 Delivering EaaS Through Cloud Computing***

Cloud computing has augmented the delivery of virtual learning environments and has embellished the concept of education-as-a-service (EaaS). EaaS means that learners and institutions are increasingly able to pursue specific areas of study, without interruption of full programs and degrees. Educational thinking leaders are advising higher education institutions to adopt EaaS as a delivery model to meet the needs of today’s learners [16].

Cloud computing allows users to use the entire Internet, anywhere with software, infrastructure, templates, and platforms. In education, cloud computing helps learners, teachers, and administrators alike. Cloud computing allows learners to fill their vacancies, wherever there is an Internet connection, allows teachers to instantly



upload learning materials, and allows administrators to easily communicate with each other and on storage and also saving money. [17]. Cloud computing also provides many services to learners and teachers through the use of IaaS because IaaS is a virtual system that includes computational resources, storage, virtualized infrastructure, and networking services [18]. Benefits of using cloud for delivering EaaS [19]:

- **Accessing online course delivery:** Permanently in massive open online courses (MOOCs) are using the cloud-based basic template to host their online learning platform and measure thousands of possible users. Cloud solutions also offer flexibility to access phones and tablets from multiple devices.
- **Online tasks and tests:** Teachers and learners use cloud-based solutions to support, access, complete, and submit tasks. Cloud-based mobile solutions allow educators and learners to share their thoughts, questions, and reviews with devices attached to the cloud platform.
- **Project in groups:** The cloud-based collaborative solution allows multiple users to simultaneously work and edit documents in group projects and enables seamless sharing and communication of ideas. Cloud-based email and audio/video support solutions are finding many recipients, especially with a large user base of faculty and students.
- **Reporting requirements:** Organizations need to gather data from many sources and gain meaningful understanding from such data for both regulatory and business needs. Cloud-based solutions can be used for storing, high-performance computing, and dash-boarding needs.
- **Real life simulators:** Real-life simulation applications that enhance the educational experience can be hosted in the cloud and run on demand. This reduces the total cost of ownership of the basic test, as these applications only run 2–3 months in a year (when classes are held) [20].

## 4 Conclusion and Future Directions

Cloud computing is a model of enabling on-demand network access to a common pool of optimized computing resources that can be provided and released quickly with minimal management effort or service provider interaction. With the use of applications, cloud computing is rising quickly in almost all areas, especially in the education system. Cloud computing offers learners and administrators the opportunity to access different applications and resources through the use of Web pages easily, at reduced costs, and quickly. Cloud computing in an e-learning platform that allows the incorporation of different e-learning standards to improve the interactivity of education objects. The concept of e-learning with cloud computing provides a cost-effective solution for educational organizations for their teachers and students. E-learning is the new model that provides digital content and an education-oriented environment for teachers and students using Internet technology. Smart learning is a model of learning that has been adopted for new generations of digital natives. Similarly, this paper has discussed the concept of cloud computing for e-learning

and smart learning. This paper has also presented the concept of delivering EaaS through the use of cloud computing and discusses the benefits of EaaS.

## References

1. Selviandro N, Hasibuan ZA (2013) Cloud-based e-learning: a proposed model and benefits by using e-learning based on cloud computing for educational institution. In: Information and Communication Technology-EurAsia Conference. Springer, Heidelberg
2. Chana I, Kaur T (2013) Delivering IT as a utility—a systematic review. arXiv preprint [arXiv:1306.1639](https://arxiv.org/abs/1306.1639)
3. Srinivasan A, Qadir A, Vijayakumar V (2015) Hybrid cloud for the educational sector. *Proc Comput Sci* 50:37–41
4. Radenkovic B, Despotovic-Zrakic M, Bogdanovic Z, Vujan V, Barac D (2014) Harnessing cloud computing infrastructure for e-learning services. *Facta Univ Ser: Electronics Energetics* 27(3):339–357
5. Kaur T, Chana I (2015) Energy efficiency techniques in cloud computing: a survey and taxonomy. *ACM Comput Surv (CSUR)* 48(2):1–46
6. Chris Coleman. How To Use Education-as-a-Service “EaaS” For A Tech Economy <https://woz-u.com/blog/how-to-leverage-education-as-a-service-eaas-and-tap-into-the-tech-economy/>. Last accessed on 2019/02/01
7. Kaur T (2019) Comprehending cloud computing for realizing energy efficiency. *J Gujarat Res Society* 21(6):532–541
8. <https://www.indiaeducation.net/online-education/articles/10-benefits-of-online-learning.html>. Last accessed 2021/04/21
9. Gamalel-Din SA (2010) Smart e-Learning: a greater perspective; from the fourth to the fifth generation e-Learning. *Egyptian Informatics J* 11(1):39–48
10. Kaur T, Chana I (2018) GreenSched: an intelligent energy-aware scheduling for deadline-and-budget constrained cloud tasks. *Simul Model Pract Theor* 83:55–83
11. Kaur T (2019) Energy efficiency trends in cloud computing. *J Gujarat Res Society* 21(6):542–553
12. Almajalid R (2017) A survey on the adoption of cloud computing in the education sector. arXiv preprint [arXiv:1706.01136](https://arxiv.org/abs/1706.01136)
13. Bosamia M, Patel A (2016) An overview of cloud computing for elearning with its key benefits. *Int J Information Sci Techniques (IJIST)* 6
14. Kaur T (2019) Challenging the cloud. *J Gujarat Res Society* 21:11
15. Mell P, Grance T (2011) The NIST definition of cloud computing
16. Bulla C, Basavaraj H, Sankalp M (2016) Adoption of cloud computing in education system: a survey. *Int J Eng Sci* 6375
17. Kaur T, Chana I (2016) Energy conscious allocation and scheduling of tasks in ICT cloud paradigm. In: Proceedings of international conference on ICT for sustainable development. Springer, Singapore, pp 589–601
18. <https://www.techfunnel.com/information-technology/infrastructure-as-a-service-iaas/>. Last accessed 2021/06/04
19. Zhu ZT, Yu MH, Riezebos P (2016) A research framework of smart education. *Smart Learn Environ* 3(1):1–17
20. Kaur T, Chana I (2016) Energy aware scheduling of deadline-constrained tasks in cloud computing. *Clust Comput* 19(2):679–698

# Blockchain and Cloud Technology: Leading the ICT Innovations



Shilpa and Tarandeep Kaur

**Abstract** Information and communication sector (ICT) is persistently transforming the manner of computing and provisioning of the Information Technology (IT) services and applications. Subsequently, the complexity, niceties and intricacies of the prevalent internet services and applications are also growing as well as changing. Eventually, cloud computing has emerged as one of the leading ICT innovations that has gained momentous importance owing to its benefits. Currently, blockchain technology is earning utmost prominence in the IT industry. The efforts are being initiated to merge the capabilities of both the cloud computing and blockchain technologies and yield utmost benefits from a collaborative work on these technologies. This paper attempts to lists the backend details for the technologies and discusses the benefits of them.

**Keywords** Blockchain · Bitcoin · Cloud computing · Cryptocurrency · Security

## 1 Introduction

ICT sector has observed stupendous developments owing to the affluent, convoluted and large number of services and applications being handled. With the advancements in the manner and frequency of IT services and applications, the demands for developing technologies that can very well handle such complexities and intricacies associated with the ever-changing ICT sector has grown [1, 2]. Cloud computing and blockchain technology have emerged as huge technological innovations that have envisaged good reputation in properly handling ICT services and application developments, provisioning as well as management. It is important to study, explore and analyze the impact of these technological innovations. Figure 1 demonstrates the current ICT services and application requirements owing to the changing market demands, user-driven requirements, environmental concerns and so on [3].

---

Shilpa · T. Kaur (✉)  
Lovely Professional University, Phagwara, Punjab, India  
e-mail: [shilpa.21819@lpu.co.in](mailto:shilpa.21819@lpu.co.in); [tarandeep.24836@lpu.co.in](mailto:tarandeep.24836@lpu.co.in)

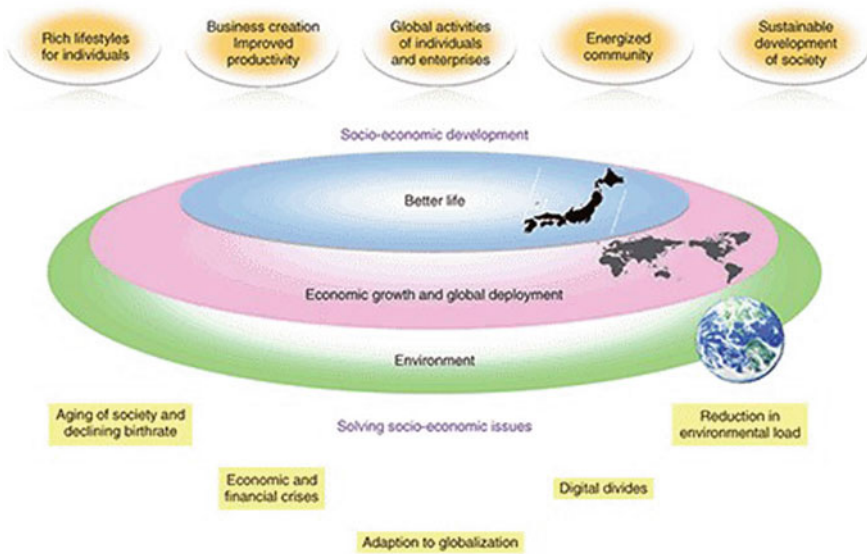


Fig. 1 Current ICT innovations requirements [3]

## 2 Blockchain and Cloud Technology

The current ICT innovations are led by cloud computing and blockchain technologies. The provisioning of IT services and applications using the capabilities of both are in prominence. Leading IT firms, such as, Google, Amazon, Microsoft have adapted and explored the potential of cloud computing and blockchain technology.

### 2.1 Blockchain Technology

Blockchain is a disseminated ledger technological innovation that has grown from vast amount of internet and actually represents a jiffy segment of the internet. However, the former internet era democratized informational exchanges while blockchain assures democratizing the real value exchanges. Blockchain was initially termed by Satoshi Nakamoto, who had delivered a protocol “A Peer-to Peer Electronic Cash System” and produced digital currency called as Bitcoin that was formed on blockchain technology [4, 5].

Blockchain acts as public transaction that can be accessed by each and everyone but there is no any central authority having control on blockchain. It enables the various sophisticated technology for professionals as well as for companies to collaborate with trust and transparency. There are various applications that are related with blockchain but the best-known applications of blockchains are the cryptographic

currencies such as Bitcoin, but many other applications are also possible. Blockchain technology is examined as driving force of the next fundamental revolution in information technology [6–8].

Cryptocurrencies are varied from the contemporary fiat money as they are not issued by national state. They are managed in a globally accessible ledger or spreadsheet listing all the transactions leveraging the resources of large P2P network for reviewing and approving every bitcoin transaction. However, blockchain is not just cryptocurrencies. It spans multiple applications and has significant commercial potential being explored by the IT [9].

### 2.1.1 Properties of Blockchain

Blockchain has emerged as prime transformational and disruptive technological innovations owing to following properties [10].

- *Dispersed*: Blockchain follows a protocol that involves distributed mathematical computations to maintain data integrity while exchanges are being done amongst variable number of computational devices irrespective of third-party involvements. There is no centralized database used in blockchain that can be compromised, instead, blocks of blockchain are globally running on individual computers.
- *Encrypted System*: In order to infer virtualized security mechanisms, blockchain makes use of 2-key authentication system.
- *Inclusive and transparent*: Being public blockchain, there exists transparency in transactions as anybody can view, download and use the transactions on a network. It is actually “a simplified payment verification mode” which can be accessed on the mobiles as well, thus eliminating the need for any kind of documentation or corporeal proofs that can be otherwise compromised.
- *Immutable*: The transactions can be readily verified, cleared and stored block wise while linking to the previous blocks for creating the chain. This linking of blocks with previous blocks ensures validity. Additionally, this helps in assuring permanent timestamps and storing of value-based exchanges, consequently, preventing ledger alterations of any kind.
- *Historical*: Since block chain is a distributed ledger representing a network consensus of every transaction that has ever occurred, the block chain must be preserved in its entirety. Storage matters a great deal. It allows the provenance of products, art, and diamonds to be traced easily.

The value of this new technology is that it allows trusted transactions to occur directly between two or more total strangers, authenticated by mass collaboration on a network of interlinked devices, and motivated by collective self-interest, rather than by profit motivated corporations or governments motivated by maintaining power and may be interested in surveillance of its citizens.

### 2.1.2 Benefits of Blockchain

In today era, various fields' researchers as well as vendors are taking the advantage of blockchain technology. It is considered as secure method in every field such as business sector, industry, healthcare sectors as well as in finance. Blockchain technology is considered as decentralized and trustless as a result it leads to new opportunities [4, 5, 10]. The following are the benefits of blockchain technology:

- *Decentralization*: Blockchain act as ledger which is not stored in single location means not centralized. There is no single authority who can handle. Due to decentralization, it can be easily verified as well as can be modified. Here, the servers and clouds are globally interchangeable with the nodes. As a result, it supports easier verifications with no third-party interactions or interferences.
- *Transparency*: Blockchain ledger is authenticated and accessible for certain users or participants and only they have the capability to make changes or perform any modifications in the information contained. Thus, the blockchain data possesses consistency, accuracy and transparency.
- *Security*: Upon the transaction approval, encryption and linking of transactions is done with the earlier transactions. This eliminates compromises for the intruders to attack on the information. Blockchain opportunistically adds complexity to the information to make it more secure from any fraudulent and unauthorized accesses.
- *Speed*: Due to non-payment cycle availability, the processing of transactions is faster and easier, thus yielding speedy activities as well as reliability.

### 2.1.3 Application Areas of Blockchain Technology

Owing to the vast security and transparency, blockchain technology is being used by various sectors which are:

*Financial Sectors*: In financial fields, blockchain technology has taken its own space due to security, transparency and simplicity, hence, finance business taking the benefit of this.

*Government*: Due to security blockchain technology is used to keep the records secretly as well for authentic purpose Such as voting records.

*Healthcare*: Blockchain technology is widely used in health sectors also. Hospital's authority used this technology to keep the record of patient authentic as well as private. Hence, it cannot be used by third-party without approval of authority.

*Educations*: Nowadays, mostly education institutes' even universities are also using the blockchain technology to keep the record of document private as well as to process them easily.

## 2.2 Cloud Computing

Cloud computing has fulfilled the long-lived dream of proffering utility as a service. It has significantly contributed in transformations of Information Technology (IT) service provisioning and shaping the manner in which services are consumed and utilized. With the origin of cloud, the innovators and developers are no longer bothered of making any capital investments in hardware in order to deploy their services or gain expertise in operating them. Exclusively, the elasticity and potential offered by cloud is unprecedented in the history of IT [1, 2, 11–15].

Cloud computing has revolutionized the provisioning and consumption of the information and computing by the organizations. The National Institute of Standards and Technology defines cloud computing as a model that enables opportunistic, accessible, vast connections to large and united pool of computational resources that can be quickly provisioned and freed with nominal service provider involvement or managerial efforts [1]. It includes provisioning of IT-relevant facilities and capabilities “as a service,” thus letting the users to entrée technical services from the cloud even without knowledge, experience or organizational control over the technical infrastructure supporting them [15]. Thus, cloud can be considered as an assortment of hardware, software and other resources accessible over Internet and provide an accumulation of solutions on-demand [1].

Today, data centers, acting as “Cloud providers,” are affianced into provisioning of numerous sprightly and valuable services to the users. The data center facilitates the users to locally or remotely access the services offered by them and by creating virtual instances instantly. A large number of IT companies such as Amazon, Yahoo, Google, Oracle, Facebook and Microsoft have setup their own data centers for providing utility-oriented cloud services [16].

Cloud computing espouses agile, reliable, thrift, scalable service delivery and computations benefitting both the consumers as well as the providers. Economically, the consumers are required to pay only for the resources required and used by them whilst the providers are liberated of cost incurred by the identifying and freeing under-utilized resources. However, the utmost priority of the service providers lies in maximization of the profits [17]. Presently, the data centers acting as “Cloud providers” are affianced in facilitating nimble and effectual ICT services and applications to the users. Cloud data centers (CDCs) are incessantly expanding and have made it possible for the users to instantaneously create virtual instances for delivering ICT services [16, 17].

### 2.2.1 Benefits of Cloud Computing

Cloud computing is mammoth of numerous benefits that can satiate any organization with its IT needs. It offers distinguishing benefits including [1, 2, 11–13, 18, 19].

*Multitenancy:* Multitenancy implies the services offered by diverse cloud service provisionary being carried out simultaneously while being located and operated in

a single data center. It makes sharing of resources and associated costs available to large number of users, thus aiding centralized infrastructure allocations that too at lesser costs; better utilization and system efficiency improvements.

*Globally Distributed:* Cloud data centers, these days, are globally distributed and located at multiple locations. Different users possessing internet connectivity on their devices can easily access cloud services irrespective of their locations from these data centers. In other words, the device and location independence enable the users to access services regardless of their location or what device they are using. Comparatively, a service provider can benefit from the geo-diversity spread across multiple geographical domains to maximize the service utility.

*Economical:* From a user perspective, cloud computing offers independence from hosting enormous and expensive infrastructure that could have been otherwise needed to access the IT services. Instead, cloud data centers offering services house huge infrastructure, monitored and maintained around the clock by service providers. The availability of vigorous infrastructure typically supported by a third-party and not acquired by the users lowers the incurrence costs.

*Resource Pooling and Sharing:* The large pool of resources is available with cloud infrastructure where resources are dynamically allocated to the multiple cloud users on-demand. The dynamic allocation and provisioning of resources enables flexibility in management and organization of resource utility levels and associated operating expenditure.

*Scalable:* The benefits associated with the dynamic and on-demand allocation & deallocation of resources on cloud are incomparable to the traditional model where the resources are provisioned driven by the peak demand. The dynamic (“on-demand”) resource provisioning aids in lowering the operational costs. Moreover, such dynamic provisioning favors scalability and elasticity.

*Pay-per Use Model:* Cloud computing applies and implements the model using pay-per-use pricing scheme. Even though this scheme for incurring pricing is dependent on the type of service but on a whole, it lowers the service operating cost.

### 2.2.2 Cloud Computing Applications and Service Providers

Cloud computing offers distributed; agile; transparent; seamless integration and delivery of services to the users and thus extends a wide horizon of opportunities for supporting highly pervasive technological progress and innovations. Moreover, by facilitating technological advancements that empower interactions across the globe, cloud computing generously proffers ample and flawless delivery of large number of services and applications [20].

The capability to host large number of applications that can be delivered to the users speedily at minimal cost has augmented and attracted cloud computing as a favorable paradigm for the IT industry [21]. The cloud services and applications are distributed across multiple domains and are accessible to the users without any location confinements. Consequently, big IT giants such as Amazon, Google, Microsoft,



Salesforce, and Facebook have significantly adopted cloud computing as principal technological platform for their operation. Owing to the extensive implementation of virtualized servers, consolidation and useful service provisioning, cloud computing has unexpectedly revolutionized and dominated the business and service delivery patterns of the IT applications. The prominent ventures include:

**Amazon Web Services (AWS):** AWS is a platform supporting design and development of highly flexible applications that can sustain scalability, storage; messaging across elastic infrastructure vested by Amazon. The prominent compute services include: Amazon Elastic Compute Cloud (EC2), Amazon Elastic MapReduce, etc. Similarly, the storage services include: Amazon Simple Storage Service (S3), Amazon SimpleDB, etc. Other applications include: Amazon CloudWatch, Amazon Simple Storage Service (SES), etc., [22].

**Google AppEngine:** IT bigwigs like Google largely offers PaaS through its venture, Google AppEngine. It provides services for development and hosting of web applications on the distributed, dynamic and scalable infrastructure of Google. The application developers can develop their applications in Java, Python and Go languages. It constantly monitors the application's usage of resources and services [23].

**Microsoft Windows Azure Platform Appliance:** The Windows Azure platform is deployable as an appliance on third-party data centers and constitutes the infrastructure controlling the physical machines in the data center. The prime Windows Azure platform appliance products include Windows Azure, SQL Azure and Microsoft-assisted network; storage; and physical configurations [24].

**Salesforce.com:** Salesforce is primarily a Software-as-a-Service (SaaS) solution providing support for Customer Relationship Manager (CRM) applications [25].

### 3 Conclusion

The ICT sector has made significant progress owing to rich, intricate service delivery patterns. Cloud computing and blockchain technology are the currently progressive faces of the ICT industry. Blockchain primarily involves decentralized databases. Decentralization is a technique that is used everywhere rather than computers and network. It involves businesses and people for the purpose of authentication and privacy of the data. However, the entire decentralization may be very hard to accomplish and it might take time for the blockchain to turn out to be in reality decentralized. Cloud computing also has revolutionized the ICT sector. Both the technologies are the primary backbone technological innovations of the IT sector in the global domain.

## References

1. Kaur T, Chana I (2015) Energy efficiency techniques in cloud computing—a survey and taxonomy. *ACM Comput Surv* 48(2). <http://dx.doi.org/10.1145/2742488>
2. Kaur T, Chana I (2017) GreenSched: an intelligent energy aware scheduling for deadline-and-budget constrained cloud tasks. *Simulation Modeling Pract Theor* 82:55–83. <https://doi.org/10.1016/j.simpat.2017.11.008>
3. ICT innovations, Available at: <https://www.nttreview.jp/archive/ntttechnical.php?contents=ntr201005sf1.html>
4. Blockchain Technology: What it means. Available at: <https://blogs.iadb.org/caribbean-devtrends/en/blockchain-technology-explained-and-what-it-could-mean-for-the-caribbean/>
5. Blockchain Tutorial. Available at: <https://www.javatpoint.com/blockchain-tutorial>
6. Ahram T, Sargolzaei A, Sargolzaei S, Daniels J, Amaba B (2017) Blockchain technology innovations. In: 2017 IEEE Technology & Engineering Management Conference (TEMSCON), Santa Clara, CA, pp 137–141
7. Kshetri N (2017) Can blockchain strengthen the Internet of Things? *IT Professional* 19(4):68–72
8. Samaniego M, Deters R (2016) Blockchain as a service for IoT. In: International conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Chengdu, pp 433–436. <https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.102>
9. Bonneau J, Miller A, Clark J, Narayanan A, Kroll JA, Felten EW (2015) SoK: research perspectives and challenges for bitcoin and cryptocurrencies. In: 2015 IEEE symposium on security and privacy, pp 104–121
10. Jesse Y-H, Ko D, Choi S, Park S, Smolander K (2016) Where is current research on blockchain technology?—a systematic review. *PloS One* 11(10)
11. Kaur T, Chana I (2014) Energy efficient cloud: trends, challenges and future directions. In: International conference on next generation computing and communication technologies (ICNGCCT 14)
12. Kaur T, Chana I (2016) Energy aware scheduling of deadline-constrained cloud tasks. *Clust Comput* 19(2):679. <https://doi.org/10.1007/s10586-016-0566-9>
13. Chana I, Kaur T (2013) Delivering IT as a utility—a systematic review. *Int J Foundations Comput Sci Technol* 3(3). <https://doi.org/10.5121/ijfcest.2013.3302>
14. Ajith Singh N, Hemalatha M (2012) Cloud Computing for academic environment. *Int J Information Commun Technol Res* 2(2):97–101
15. Mirashe SP, Kalyankar NV (2010) Cloud Computing. *Commun ACM* 51(7):9
16. Teng F (2011) Management of data and scheduling of tasks on architecture distributes. PhD Thesis, École Centrale: A University Institution, Paris
17. Lee YC, Zomaya AY (2012) Energy efficient utilization of resources in Cloud computing systems. *J Supercomputing* 60(2):268–280
18. Beloglazov A, Abawajy J, Buyya R (2012) Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. *Futur Gener Comput Syst* 28:755–768
19. Garg SK, Yeo CS, Anand asivam A, Buyya R (2011) Environment-conscious scheduling of HPC applications on distributed Cloud-oriented data centers. *J Parallel Distributed Comput* 71:732–749
20. Fernandes RT. Green-cloud: economics-inspired scheduling, energy and resource management in cloud infrastructures. <https://www.inescid.pt/publications/10630/pdf/>
21. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of things (iot): a vision, architectural elements, and future directions. *Futur Gener Comput Syst* 29(7):1645–1660
22. Amazon web services. Available on: [www.aws.amazon.com/](http://www.aws.amazon.com/). Last accessed August 2016
23. Google app engine. Available on: [www.cloud.google.com/appengine/](http://www.cloud.google.com/appengine/). Last accessed June 2016
24. Microsoft azure. Available on: [www.azure.microsoft.com/en-in/](http://www.azure.microsoft.com/en-in/). Last accessed July 2015
25. Salesforce.com. Available on: [www.salesforce.com/in/](http://www.salesforce.com/in/). Last accessed November 2017

# Automated Gate System Using Number Plate Recognition (NPR)



Shreya Anekar , Samrudhi Yeginwar , and Harsha Sonune

**Abstract** The utilization of vehicles in our day-to-day existence is rising dramatically, and as expanding vehicles are violating the principles of traffic, robbery of vehicles, ingoing to restricted areas, and strange number of mishaps lead to increase in the crime percentages. For any vehicle to be recognized, a vehicle number plate will play a significant part in this world. It basically increases the safety and security of the system. One of the applications of the number plate recognition (NPR) system is that it can be used at entry points of various offices, societies, etc., for authentication of the vehicles. This process can be fully automated. In most of the vehicle entry points, the entry method for guests, staff, or understudies making visits to the area incorporates a security labor force who confirms the details by checking the recognizable proof archive. Since this is a manual technique, the strategy has high probability that it might cause human mistakes during the affirmation and enrollment procedure. The task of registering and composing information physically could be troublesome and tedious for the personnel appointed at the designated spot. It is difficult to share the information, as the information entered is in the printed version design. This paper proposes an automatic authentication procedure, which will cut off the manual work and limit the security individual's work. The proposed system will lessen registration time and will also have other added benefits regarding parking management and traffic monitoring.

**Keywords** Automated gate · Number plate recognition (NPR) · Image processing · Computer vision · Optical character recognition (OCR) · Convolutional neural network (CNN) · Database

---

S. Anekar (✉) · S. Yeginwar · H. Sonune  
Department of Information Technology, MKSSS's Cummins College of Engineering for Women,  
Pune, Maharashtra, India  
e-mail: [shreya.anekar@cumminscollege.in](mailto:shreya.anekar@cumminscollege.in)

S. Yeginwar  
e-mail: [samrudhi.yeginwar@cumminscollege.in](mailto:samrudhi.yeginwar@cumminscollege.in)

H. Sonune  
e-mail: [harsha.khedkar@cumminscollege.in](mailto:harsha.khedkar@cumminscollege.in)

## 1 Introduction

At any designated spots, in the present scenario, a rider needs to step outside from the vehicle and give his vehicle's details which are enrolled physically in a register. This interaction between the officials who are keeping up the evidence at the entryway and the visitors turns repetitive as well as tedious as they need to stand in the queue just for the little information. Additionally, maintaining the records and the physical register may lead to human mistake, which may prove incorrect when in a crisis while giving the details, subsequently not having the option to follow the legitimate individual. If in a case, the system crashes or goes disconnected due to the power cutoff and natural disasters, the records maintained in the excel document or register will not guarantee the availability of data, and we will not have the option to get the required data [1]. Thus, an automated gate system is a need of an hour.

The framework aims at identifying the number plate of vehicles using image processing which will lessen the work of the security official as well as save the time of the visitors. A challenge that these systems face is that they have a lot of constraints like detecting number plates in less lighting is difficult, designer number plate detection is quite challenging, and also, weather condition is a major factor [2].

In India, there are different formats for number plates with varying fonts, color, size, etc., [3]. Due to the large population and heavy traffic, this automated system is the need which will help in saving the time and fuel. Thus, the automated system should have less constraints and fast detection speed to ease the work.

This automated gate system using number plate recognition will wipe out the manual entry procedure and limit the security individual exertion at different spots like big societies, institutions, workplaces, offices, etc. Using this system, security and approved personnel (admin) can track vehicles alongside its owner's data. The number plate recognition (NPR) system will automatically detect the number plate using image processing, when a vehicle comes in the entry point of premise and will permit the vehicle after verification [4]. The system also monitors the movement of vehicles along with their owner details, and these details can be accessed at anytime from anywhere if permitted by the authorities. The association of this paper is portrayed as follows: Sect. 2 sheds some light on previous work done in this type of projects. Proposed system is clarified in Sect. 3. At last, Sect. 4 gives a conclusion and future scope.

## 2 Related Work

There are various researches and methods for number plate recognition but limited work available in the fields of automated gates.

Kashyap et al. [5] present a calculation for number plate location and letter recognition. The calculation utilizes optical character recognition (OCR) which coordinates the numbers and characters of the number plate with one another which

is considered as the main stage. This idea of automatic number plate recognition (ANPR) depends on the template matching algorithm and achieves 75–85% accuracy for Indian number plates.

Khinchí [6] presents an overview of an automatic number plate recognition technology and various methods. It presents the steps that can be taken to detect the number plate.

The paper [7] presents a prototype of automated gate using NPR which operates using Raspberry Pi microcontroller, Node-RED software is used for authentication, and OpenALPR library is used for image processing. The accuracy for image taken from specific height and distance is 87.50%.

In paper [1], they have developed a Web-based application using a waterfall model for an automated gate system. This paper gives us a brief about steps to be followed for creation of NPR system and also creation of database. The Roberts vertical edge detection is utilized for edge detection purposes.

Kulkarni et al. [3] recommend a methodology explicitly for Indian number plates. India has various kinds of number plates. The framework comprises a blend of calculations like ‘component-based number plate localization’ for finding the number plate and ‘image scissoring’ for character division. The structure can perceive single-line and double-line license plates under various enlightenment conditions with a success percentage of about 82%.

Distinctive number plate recognition (NPR) procedures have been presented over the most recent couple of years and are ordered dependent on the highlights, such as edge detection, image capturing, surface, shading, and also, character highlights they use.

Table 1 gives the comparison of various number plate detection systems using their accuracy. The accuracy of each system changes depending upon the number of characters present in the number plate [8].

**Table 1** Comparison of NPR techniques

Sr. No	Technique used	Number of characters in the number plate	Average accuracy (in %)
1	BPNN [9]	8	96
2	Morphological [10]	9	92
3	CNN [11]	9	94
4	CNN	6	92
5	Localization [12]	8	78
6	OCR [13]	7	95
7	OCR [5]	10	80

### 3 Proposed System

#### 3.1 System Architecture

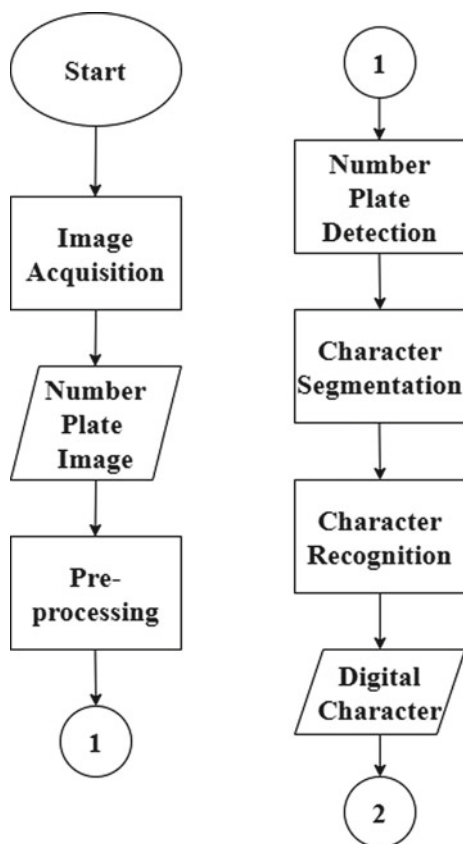
The entire framework is partitioned into two sections. The initial stage of the system is number plate recognition (NPR) using image processing, and the next stage is the authentication of detected number plates.

The flowchart in the figures below depicts the system architecture. Figure 1 represents the recognition of number plate, and Fig. 2 represents the authentication of detected number plates.

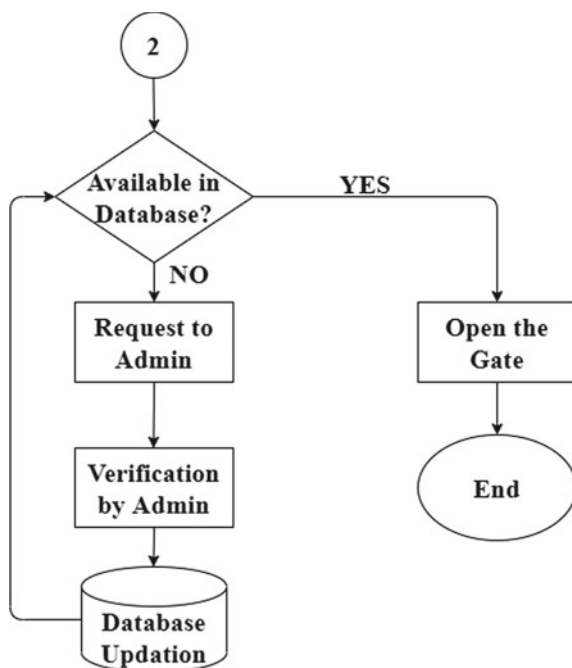
The proposed system has the following advances:

1. Image acquisition: Picture of vehicle is captured.
2. Number plate image: The captured vehicle image is displayed.
3. Preprocessing: Obtaining an expected picture and eliminating the bends.

**Fig. 1** Part-I: Number plate recognition (NPR) system



**Fig. 2** Part-II:  
Authentication of detected  
number plate



4. Extraction of number plate: Eliminate insignificant parts and concentrate the number plate.
5. Character segmentation: Segmenting or dividing the characters from the detected number plate.
6. Character recognition: Perceive letters and show the obtained results.
7. Authentication:

Is detected number plate available in the database?

A.YES: Open the gate.

B.NO: Request is sent to admin.

8. End

### 3.2 Number Plate Recognition

Number plate recognition involves the following steps:

- Image acquisition—Image acquisition implies catching a picture from any computerized camera or portable camera. Pictures are taken in various edges and light conditions. Distance from the image is taken, and atmospheric conditions (winter, brown haze, precipitation, and summer) may cause the ‘noise’ at

the hour of capturing photos. The images captured are colored images which are in RGB format in terms of image processing.

- **Preprocessing**—The essential objective of preprocessing is to improve the difference of the input picture, to diminish the noise in the picture, thus to upgrade the preparing speed, and to expand the perceivability and nature of input images. In the proposed approach for NPR, in the pre-handling stage from the outset, the RGB image is changed over into grayscale image, and at that point, filters can be utilized to eliminate the noise from the grayscale image.
- **Image binarization**: The image is changed over to high contrast design in this progression. The reason for applying shading change is to diminish the quantity of scopes of the shading scale from (0–255) to (0–1). In this progression, the deducted grayscale image is changed over into a binary image.
- **Edge detection**—Different edge recognition calculations are accessible to utilize like Sobel edge identification calculation and Canny edge location calculation. They can detect edges also with noise suppressed. These operators are the most easy and popular way to perform edge detection.
- **Extracted plate region enhancement**—Region is extricated from the picture after the discovery of number plate region. The separated number plate may have distinctive noise, unrequired openings, outlines, nut bolt, and so on. Consequently, improvement of number plate is done.
- **Character segmentation and character recognition**—Character segmentation (CS) carries on as a scaffold between the number plate extraction and character acknowledgment part. In this progression, the characters on the number plate region are isolated or fragmented. There are numerous elements, for example, picture noise, casing of plate, revolution of plate and brightening distinction, and so forth that make the character division hard. Character acknowledgment (CA) is the last step of the NPR framework. The contributions to this progression are divided characters, and result of this progression is recognized number from number plates.

### 3.3 Authentication

The result of picture processing will be a computerized character that can be included in the data set. On the off chance that the data set as of now has that particular data, its timestamp can be added to the current data, or another segment in the data set can be added if the data coming is new. Data that has been put away in the data set can be seen by a security official with his phone or PC.

Table 2 shows the theoretical perspective on the vehicle number information portrayal in the data set with entry time and owners' information.

The developed system and the database of the owners should have high-level security so that no third member other than admin can use this confidential information.



**Table 2** Abstract view of database

Sr. No	Vehicle number	Date	Entry time	Owner details
1	MH34AM6876	03/08/2020	1:00 PM	Akshaya
2	MH16BH2165	15/12/2020	10:30 AM	Shreya
3	GJ12AK4447	12/02/2021	3:00 AM	Samrudhi
4	TN09SI2316	25/07/2021	3:30 AM	Shrutika
5	MH16BH5130	30/09/2021	11:26 PM	Janhvi

Finally, as per the requirement of the framework, the actual plan can be modified and data can be updated.

4 Conclusion and Future Scope

An automated gate system at entryway focuses basically to restrict the burdens of safety officers and the guests. It is executed by utilizing NPR framework which extracts the number plate of vehicles from the captured pictures and converts into text format. After the recognition of the individual character, these characters are included in a data set with its particular timestamp of entry and exit, which is made open to a concerned official with the Web application assistance by means of PC or cell phone having Internet access to make any necessary moves inside advantageous time. The system can be used by safety officials to track the vehicle’s movement passing the entry points and that information may be gotten to by officials for any necessary activities. Moreover, they will actually want to check the proprietor’s information with the assistance of the vehicle’s number plate anyway the right proprietor cannot be ensured by the framework on account of no facial acknowledgment carried out in the framework. The execution of face recognition will be the future extent of this exploration. The framework should work out adequately for wide assortments in light conditions and different sorts of number plates regularly found in India. In India, it is certainly a superior option in contrast to the current manual system. As of now, there are sure constraints like speed of the vehicle, slant in the image, and script on the number plate which can be fittingly disposed of by upgrading the calculations further.

References

1.

Tenzin S, Dorji P, Subba B, Tobgay T (2020) Smart check-in check-out system for vehicles using automatic number plate recognition. In: 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)

2.

Kaur S (2016) An automatic number plate recognition system under image processing. Int J Intelligent Syst Appl

3. Kulkarni P, Khatri A, Banga P, Shah K (2019) Automatic number plate recognition (ANPR) system for Indian conditions. In: 19th International conference Radioelektronika. IEEE
4. Chowdhary N, Kumar A, Singla S (2019) Image processing based intelligent parking system using number plate recognition. *Int J Sci Technol Res* 8(11)
5. Kashyap A, Suresh B, Patil A, Sharma S, Jaiswal A (2018) Automatic number plate recognition. In: 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), IEEE, pp 838–843
6. Khinchi M, Agarwal C (2019) A review on automatic number plate recognition technology and methods. In: International Conference on Intelligent Sustainable Systems (ICISS). IEEE, pp 363–366
7. Al-Mahbashi LTA, Yusof NAT, Shaharum S, Karim MSA, Faudzi AFM (2019) Development of automated gate using automatic license plate recognition system. In: Proceedings of the 10th national technical seminar on underwater system technology 2018. Springer, Heidelberg, pp 459–466
8. Kumar A, Verma D (2020) Number plate reorganization using image processing and machine learning approaches: a review. In: 2020 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore
9. Barnouti NH, Naser MAS, Al-Dabbagh SSM (2017) Automatic Iraqi license plate recognition system using Back Propagation Neural Network (BPNN). In: Annual conference on New Trends in Information & Communications Technology Applications (NTICT), IEEE
10. Prabhakar P, Anupama P, Resmi SR (2014) Automatic vehicle number plate detection and recognition. In: International conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE, pp 185–190
11. Musaddid AT, Bejo A, Hidayat R (2019) Improvement of character segmentation for indonesian license plate recognition algorithm using CNN. In: International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)
12. Rana N, Dahiya PK (2017) Localization techniques in ANPR systems: a-state-of-art. *Int J Adv Res Comput Sci Softw Eng* 7(5)
13. Sharma G (2018) Performance analysis of vehicle number plate recognition system using template matching techniques. *J Information Technol Softw Eng* (2)

# Securing the COVID Patients' Medical Records Using Encrypted Image Steganography



Asha Durafe  and Vinod Patidar 

**Abstract** Strengthening the field of medical informatics has become extremely important and challenging in today's era of digital image transmission. The present work discusses a novel image steganography method that can withstand either statistical tool-based or machine learning (ML)-based steganalysis attacks to ensure secure transmission of the medical images like chest X-rays, CT or MRI reports to the medical experts or hospitals for early diagnosis and fast recovery of the infected persons. The Rivest–Shamir–Adalmen (RSA) encryption technique is incorporated to convert a secret medical image into a cipher image which is enveloped inside a unique fractal cover thus enhancing robustness of the proposed method without debasing perceptual quality of the image. A 1-least significant bit (LSB), 2-LSB and 4-LSB spatial domain method is implemented and compared for hiding the encrypted secret medical image inside a fractal cover image which adds another layer of security. The novelty of the proposed algorithm is the use of fractal cover image with zero skewness angle to affirm the secure location for medical image hiding. Furthermore, steganalysis checks are performed using a deep learning feature extraction model 'Xception' and 'WEKA' image classifiers which authenticate security and integrity of the patients' confidential data. Stego performance of the proposed method is also evaluated by steganalysis tools such as StegSecret, StegExpose and histogram analysis. The statistical and differential analysis performed demonstrates the efficiency of the proposed algorithm over the existing techniques.

**Keywords** Medical Image Steganography · RSA · LSB · Steganalysis · Xception · WEKA · StegSecret · StegExpose

---

A. Durafe (✉)

Shah & Anchor Kutchhi Engineering College, Chembur, Mumbai 400088, India  
e-mail: [asha.durafe@spsu.ac.in](mailto:asha.durafe@spsu.ac.in); [asha.durafe@sakec.ac.in](mailto:asha.durafe@sakec.ac.in)

A. Durafe · V. Patidar

Sir Padampat Singhanian University, Bhatewar, Udaipur, Rajasthan 313601, India  
e-mail: [vinod.patidar@spsu.ac.in](mailto:vinod.patidar@spsu.ac.in)

## 1 Introduction

During the coronavirus disease 2019 (COVID-19) outbreak, the need of ensuring and getting information turns out to be significantly more fundamental while it is sent through correspondence to medical organizations. This prerequisite turns out to be significantly more articulated if the information we are discussing is the clinical data of a patient. Information security and insurance are the major questions looked at by clinical experts. It is imperative to get the clinical information while sending it over to a specific organization or in any event, safeguarding it on a data set. COVID-19 infection influences the respiratory arrangement of people, and various sorts of medical images are a vital part of the counteraction and treatment of this sickness. These medical images incorporate X-beams and CT outputs of the chest of the influenced individual. These days, it is a pandemic circumstance where individuals are a lot stressed over social distancing, and therefore, it is compulsory to get the individual information of the patient affected and safely communicate it over the networks. This is not just for COVID-19 but has become an important essential requirement for all clinical information carrying medical images like CT, MRI, X-ray or any kind of medical informatics data. The patient subtleties ought to consistently be kept secret. This is the place where cryptography and steganography assume a significant part.

The main frontiers in the medical image hiding paradigm are image encryption, digital watermarking and image steganography [1]. By a wide margin, asymmetric encryption is the highest reliable method of cryptography [2] because of its key size flexibilities. Asymmetric cryptography tends to the issue of key distribution for encryption which remains a serious issue in symmetric cryptography. In cryptography, the size of the key length proportionally increases the security of the algorithm. This likewise brings a significant benefit to asymmetric cryptographic calculation since it has a more drawn out key length and, accordingly, makes it more robust against attackers. Relatively, speed is a significant disadvantage in asymmetric encryption because of the intricacy of its numerical calculations. There is a compromise between security and the speed in asymmetric encryption techniques. The proposed method, along these lines, utilizes asymmetric encryption because of its undeniable benefit. As much as we keep on imparting in an untrusted medium like the Web, security stays at its highest need. Symmetric encryption techniques such as block cyphers and stream cyphers are quick and productive for large payloads, but the hindrance is the need to maintain the key secret. This can be particularly difficult where scrambling and unscrambling happen in different areas, requiring the key to be moved securely between areas. Notwithstanding the tremendous measure of contributions made by chaotic encryption [3, 4], its applications and real-time implementations are complex when contrasted with traditional cryptography [5]. Keyed hash functions-based cryptography is another widely accepted method of cryptography but suffers from probable efficiency loss especially while hashing large databases [6]. The process of adding visible or invisible structures to digital images is called digital watermarking which also plays a vital role in image hiding and information

security disciplines. However, image manipulation such as cropping or resizing may lead to unreadable watermarked images.

Recently, the distributing and broadcasting ventures are getting intrigued by procedures for stowing away scrambled copyright imprints and serial numbers in computerized films, sound recordings and interactive multimedia commodities. Moves by different governments to limit the accessibility of encryption techniques have persuaded individuals to study techniques by which private messages can be hidden in apparently harmless cover messages using steganographic tools [7]. So to add another layer of security against the attackers, steganography techniques with encryption are preferably implemented instead of using pure steganography without encryption.

### ***1.1 Contribution of the Work Presented in This Manuscript***

This manuscript presents a novel medical image hiding scheme using RSA asymmetric encryption and LSB (least significant bit) spatial domain embedding with formula-driven cover images which can withstand statistical as well as ML-based steganalysis attacks resulting in reliable privacy protection technique in the field of medical informatics. Although RSA encryption for digital images is slower among the several other encryption techniques, it has following advantages which cannot be ruled out:

1. Difficulty in factorization of prime numbers which causes RSA to still remain secure and complex to decrypt [8].
2. RSA encryption fixes the issues like key management and key distribution [9].

Hence, the proposed method is implemented with RSA encryption as the first layer of security. Furthermore the 1-LSB, 2-LSB and 4-LSB steganography model with unique formula-driven fractal images offer the second layer of security. The exhaustive performance analysis evinces that the proposed medical image transmission framework can impede against privacy breaches.

The stego image is analysed with the steganalysis tools StegSecret and StegExpose to demonstrate the efficiency of the algorithm. The statistical histogram steganalysis is also carried out to evaluate the spatial difference in the cover image and stego image. Furthermore, ML-based steganalysis using the Xception feature extraction model is carried out to ensure confidentiality and integrity of the proposed medical image steganography algorithm. Furthermore, the subjective and objective metrics analysis done corroborates the effectiveness of the proposed algorithm.

## **2 Related Work**

The real-time polymerase chain reaction (RT-PCR) test which is used to confirm positive coronavirus (COVID) disease has shown many false-negative results and proved

to be an insufficient tool to detect the viral RNA (ribonucleic acid) [10]. To avoid ambiguity, if any, the second and most promising approach used to detect COVID-19 virus is the medical imaging tools such as chest CT scans. The thoracic CT scans are playing a vital role to detect COVID-19 virus and complications caused due to COVID-19 such as heart failure and pulmonary or respiratory issues [11]. On similar lines, chest X-rays are also found to be helpful tools in further diagnosis of COVID-19 victims suffering from pneumonia [12]. The medical informatics field using such imaging modalities is utilizing the internet to transmit or receive the patients' private data thus necessitating the protection of it. This need can be fulfilled by steganography which endeavours to shroud the very fact of the presence of communication. Steganography with encryption aims to conceal an encrypted image inside a cover media generating a stego image that is transferred to the intended receiver where the secret image can be extracted with the help of a stego key [13].

Reyad and Karar [10] have developed a Hash Blum Blum Shub generator to encrypt the medical images of COVID-19 victims. Multiple key streams are proposed to attain a high degree of privacy required for the transmission of patient's data. Bose et al. [14] have proposed an image encryption technique using non-systematic cyclic redundancy code for encoding multiple medical images for COVID-19.

Sahu and Sahu [15] have discussed an image steganography method with dual encryption to secure X-ray scans of COVID-19 patients. This dual-encryption-based steganography scheme utilizes DNA encryption along with Baconian cyphers and LSB embedding to hide the medical images secretly. Akkasaligar and Biradar [16] have designed a medical image encryption technique using DNA encryption and a dual-hyperchaotic map which combines Taylor Chirikov map and Chen's hyperchaotic map. Paul [17] has discussed recent advances in image encryption techniques due to COVID-19. Various image encryption techniques such as chaotic encryption, DNA cryptography, selective ROI-based encryption, discrete wavelet transform (DWT)-based selective image encryption and DES encryption combined with LSB embedding are reviewed.

Zou et al. [18] have developed a steganalysis model based on deep learning and convolutional neural networks, and also, a feature learning algorithm is discussed for reduced insertion rate. Jung [19] has described various steganography and steganalysis methods using scikit-learn, TensorFlow, and Keras machine learning (ML) frameworks which are used to evaluate the stego images.

Mendez et al. [20] discussed LSB image steganography algorithm using skew Tent map and tested it using a universal steganalysis tool StegExpose. The steganalysis performed by this tool consists of three types of evaluations, viz. sample pairs test, chi-square attack test and RS steganalysis. A pixel to block conversion technique was proposed by Parah et al. [21] where the secret data was concealed using intermediate significant bit substitution instead of LSB method. However, the reported methods suffer from low perceptual quality images and offer reduced payload capacity of 0.75 bpp.

### 3 Techniques Used

#### 3.1 RSA (Rivest-Shamir-Adleman) Algorithm

It is the oldest asymmetric cryptography algorithm in which two separate keys, public key and private key, are generated for encryption and decryption, and the public key generated for the encryption process is shared publicly [22]. The data/images which are encrypted using the public key are decrypted using the private key which remains with the owner. The RSA algorithm implementation steps are enlisted below:

Key generation steps:

1. First, pick the two large prime numbers  $p$  and  $q$  in the range ( $\mp 100$ ) such that  $p \neq q$ . For security purposes, the number  $p$  and  $q$ , to be picked, should be of comparable bit length. Prime whole numbers can be productively selected by primality testing.
2. Compute key length  $n = pq$
3. Calculate  $\Phi(n) = (p - 1)(q - 1)$ . Keep  $p, q, \Phi(n)$  secret.
4. Choose encryption key  $e$  such that  $1 < e < \Phi(n)$  also  $\gcd(e, \Phi(n)) = 1$ .
5. Calculate decryption key  $d = e^{-1} \bmod (n)$
6. Thus, Public key:  $(n, e)$
7. Private key:  $d$ , where  $d$  is the inverse of  $e \bmod (\Phi(n))$ , and is calculated using the extended Euclidean algorithm.

Encryption:

1. The sender (medical expert) sends the public key  $(n, e)$  to the intended recipient keeping the private key  $d$  confidential.
2. The secret image  $I$  is converted into integers  $i$  where  $0 < i < I$  and  $\gcd(i, n) = 1$ . Calculate the cipher image  $C$  and send it to the intended recipient.

where  $C = I^e \bmod [n]$ .

Decryption:

1. Perform the inverse of step 2 in encryption and compute  $I = C^d \bmod [n]$  in which private key

$$d = 1 \bmod ((p - 1)(q - 1))$$

2. Retrieve secret image using private key  $d$ .

#### 3.2 Generation of Fractal Images

These are the self-similar images that can be created by repeatedly iterating the fractal equations. Each fractal image represents a unique image and are very sensitive to the initial conditions and parameters and therefore can be reproduced only if exact

fractal parameters are inputted. Due to these unique features of self-similarity and sensitivity, fractal images are preferred as cover images in image steganography techniques [23]. Two of the most common and very popular fractal images are the Mandelbrot set and Julia set. All the points inside a Mandelbrot set are Julia sets. These fractals can be defined with the following parameters:

Starting point: Always set to be (0, 0) for a standard set.

Power: It defines the exponent value, i.e. seed. The basic formula associated with this parameter is  $Z = Z^2 + C$  where  $C$  is the complex parameter called the seed.

Bailout value: It defines the magnitude of  $Z$  for which the formula will stop iterating. To create true Mandelbrot and Julia set fractals, its value should be chosen to 4 or greater than 4 [24].

In proposed work, the Mandelbrot fractal images are created using the Ultra Fractal 6.03 along with the unique parameter sets used for the generation. These unique fractal images are generated with zero skewness angle thus ensuring the normal distribution of colour intensities of the cover images. If the skewness angle of the cover image is greater than 3, then it loses the suitability criteria of the perfect cover image to hide the secret image [25].

### 3.3 *LSB (Least Significant Bit) Embedding*

It is the most preferred spatial domain technique which offers high embedding capacity without compromising the imperceptibility of the image [26]. In this technique, the secret image pixel data is inserted into the cover image by replacing its least significant bits with the secret image data. However, the problem associated with such techniques is that they are easily extractable by the attackers and hence can be detected with RS steganalysis. Hence, to enhance the security of the proposed method, the secret image is encrypted through the RSA algorithm before the LSB insertion is implemented, and a single bit LSB, 2-LSB and 4-LSB insertion is done so as to maintain the visual quality of stego image [27].

### 3.4 *Steganalysis*

It is the act of spotting secret content in multimedia files such as audio, images or videos. The major challenge for the digital forensic expert is to develop a general tool or technique which can be applied to all types of steganography methods. In the proposed scheme, tool-based steganalysis is performed by using StegSecret and StegExpose tools. Furthermore, steganalysis using 71 layers deep Xception model based on convolutional neural network is performed to substantiate the integrity of the proposed method.



### ***3.5 Steganalysis Using Machine Learning (ML)-Based Algorithms***

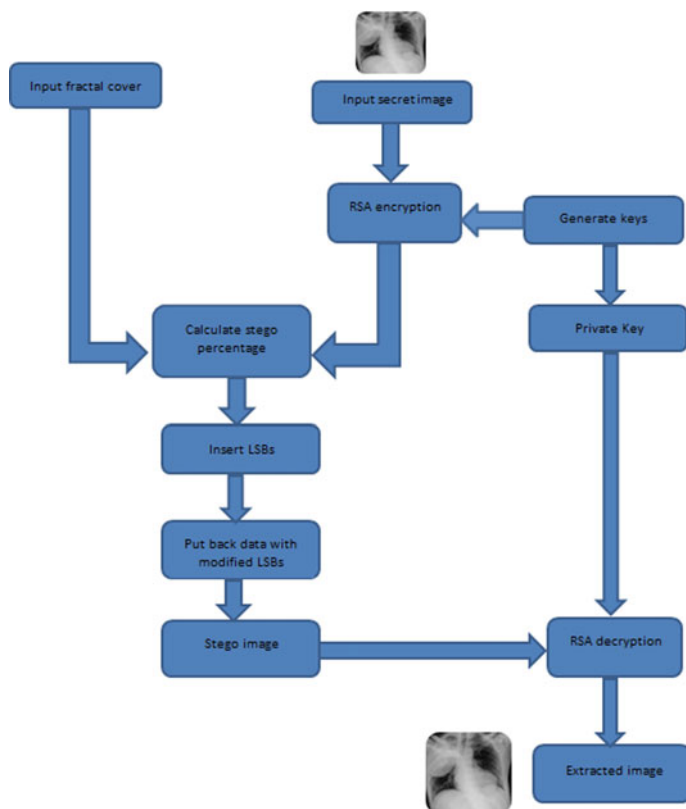
Machine learning is utilized in various types of multimedia processing such as audio, digital images, video and speech, natural language processing (NLP), computer vision and so forth since a long time. Essentially, there are numerous fields in which utilization of various types of artificial intelligence algorithms such as ML are profound including the field of steganography and steganalysis.

Image steganalysis utilizing ML is a two-stage measure to perform image classification: the initial step is to extricate the factual highlights from the embedded image and cover image and to separate the cover from stego-object. The purpose for depending on these measurable highlights is the speculation that insertion of secret image changes the factual highlights of the cover image. The second step includes the execution through ML algorithms to classify the stego images and the cover images. Subsequently, the features are extricated from the stego and cover image to train a classifier, and the test images are apportioned to the ML classifier model for examination. The lone contrast between the cutting-edge strategies is the measures picked for the feature extraction and determination [19]. Once the ML-based classifier classifies the cover and stego images with high accuracy, then the test images are evaluated to verify whether they are able to withstand the ML steganalysis attack or not. If the images created by proposed algorithm remain undetected, it indicates that the algorithm is robust enough to withstand ML detection attacks.

In the proposed method, feature extraction is done using FractalDB-60 database [28] consisting of 60,000 fractal images of various categories out of which 1500 fractal images are used as cover images. The extracted features of these cover images are classified using WEKA tool for different ML classifiers. The widely employed deep learning model 'Xception' is utilized to extract total 2047 features for each of the 1500 fractal images taken from FractalDB dataset. Another set of 1500 stego images were created using the ML-based HUGO steganography algorithm for training purpose. The testing database of another 1500 cover and stego images consisting the 24 images created by proposed algorithm is sent to WEKA which offers a collection of ML algorithms for image classification.

## **4 The Proposed Method**

A dual-layer security algorithm to hide COVID patients' secret CT images or chest X-rays has been designed using RSA encryption along with 1-LSB, 2-LSB and 4-LSB spatial domain steganography methods, and the results for each method are analysed and demonstrated. Further, steganalysis checks are carried out with state-of-the-art tools and methods which validate the proposed method. Figure 1 shows the basic stages of implementation and algorithms 1, 2 and 3 in sequence are enlisted in this section below to run the proposed system.



**Fig. 1** Proposed algorithm

---

### Algorithm 1: RSA key generation

---

**Input:** Java keypair generator

**Output:** public key, private key

1. Create a Java KeyPairGenerator instance:  
`KeyPairGenerator keyPairGenerator =  
 java.security.KeyPairGenerator.getInstance("RSA");`
  2. Initialize KeyPair generator to generate key with 1024 bits size:  
`keyPairGenerator.initialize(1024);`
  3. Call the generateKeyPair( ) method to generate a keypair:  
`keyPair = keyPairGenerator.generateKeyPair();`
- 

The Java API (application programming interface) class package is accessed by using MATLAB Java interface [29]. Java class is one of the MATLAB data types. It enables the MATLAB interface to create objects and call Java methods [30].

---

**Algorithm 2: Embedding**


---

**Input:** cover image (img), secret image (secret), public key

**Output:** stego image (stego)

1. Read the cover image.
  2. Number of nLSBs=1
  3. Read the secret image and display the size of the secret image in bytes:  
Secret\_bytes = typecast(secret, 'uint8');
  4. Get number of bytes used to store just the size information:  
Byte\_Count = uint8(numel(secret\_bytes));
  4. Encrypt secret image using RSA encryption API to get encrypted\_image.
  5. Convert encrypted secret image into 1-D array:
  6. Calculate stego percentage.
  7. Get the bytes to be modified from the Stego image.
  8. Insert LSBs.
  9. Put back data with modified LSBs to generate a stego image.
  10. Display stego image.
- 

---

**Algorithm 3: Extraction**


---

**Input:** Stego image, private key

**Output:** Extracted secret image

1. Read the stego image.
  2. Convert stego image to bits:  
data = de2bi(stego);
  3. Get LSBs:  
data = data(:, 1:nLSB); %where nLSB=1
  4. Convert into one dimension array of bits:  
data = reshape(data.', 1, []);
  5. Make data multiple of 8 for byte conversion
  6. data(end+(8-mod(numel(data), 8))) = 0;
  7. Get bytes:  
data = reshape(data, 8, []);  
data = bi2de(data);
  8. Get Header Byte Count from first byte:  
nHeaderByteCount = double(data(1));  
nEncryptedBytes = typecast(data(2:9), 'double');  
encrypted = data(10:10+nEncryptedBytes-1);
  9. Get size of secret image from subsequent bytes.
  10. Get image bytes.
  11. Reshape the bytes to get the secret image: secret bytes.
  12. Display the extracted secret image.
-

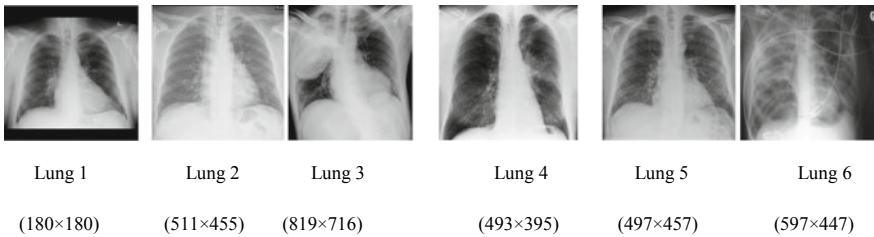
## 5 Experimental Results and Discussion:

To validate the performance of the proposed algorithm, it was simulated using MATLAB version 2021 on Intel(R) Core(TM) i3-3220 CPU @ 3.30 GHz and 4 GB RAM. The COVID affected lung images (chest X-rays) in Fig. 2 are used as secret medical images and are taken from COVID-ChestXray-Dataset [31]. The fractal cover images are of size  $2048 \times 2048$ . Figure 3 shows the fractal images (a-d) with their seed values, secret chest X-rays (e-h), corresponding stego images (a + e, b + f, c + g, d + h) and the extracted secret images. Variation in payload capacity for 1-LSB and 4-LSB methods for Lung 5 as a secret image can be seen in Fig. 4. It can be noted that 1-LSB method has low payload capacity (64% bytes of the cover image are used) but high PSNR and low MSE. However, 4-LSB method offers the highest payload capacity amongst the three methods (16% bytes of cover image are used) at the cost of low PSNR and high MSE as compared to 1-LSB method. The stego image created with Fractal 1 cover and Lung 1 secret image is assigned number 1, the stego image created with Fractal 1 cover and Lung 2 secret image is assigned number 2 and so on to create 24 stego images as seen in Table 1. Table 1 also lists the metrics PSNR, MSE, NCC and SSIM for various fractal images. Figure 5 illustrates the variation in PSNR, MSE, NCC and SSIM for all 24 stego images in Table 1. It can be observed that for all 24 tests performed, the statistical metrics are within acceptable range to offer high perceptual quality and increased security to the image.

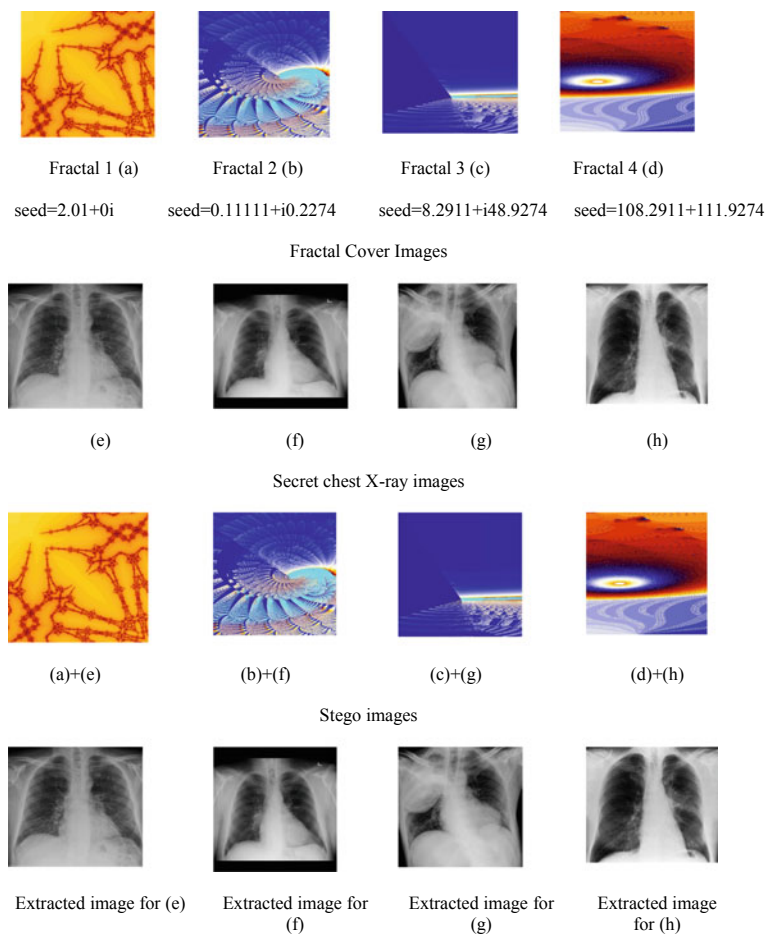
### 5.1 Statistical Steganalysis

#### 5.1.1 StegSecret Tool

It is an open-source steganalysis tool developed using JAVA JDK 1.5 multi-platform, especially targeting steganalysis of digital images/audio/video which enables the user to detect the existence of secret information. StegSecret b0.1 analyses the images against three types of attacks which are visual attack, statistical chi-square attack which can be used for sequential LSB techniques to measure the payload size and RS



**Fig. 2** Chest X-rays used as secret medical images

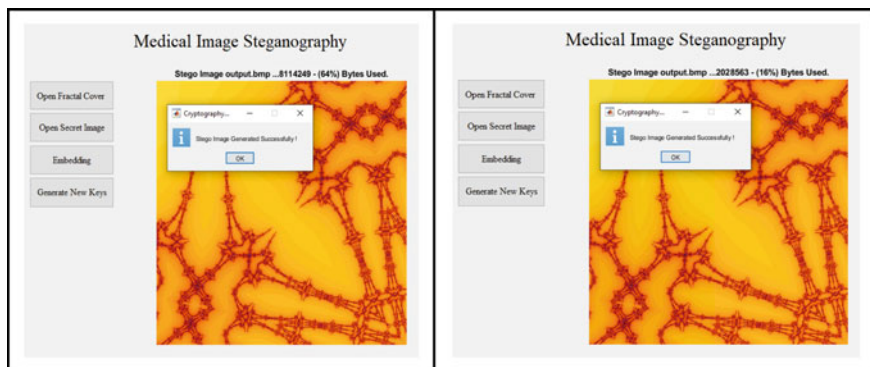


**Fig. 3** Results obtained with proposed algorithm

steganalysis attack which can also measure payload capacity and randomly scattered least significant bit recognition. StegSecret tool steganalysis was carried on total 1000 images consisting 24 stego images generated by the proposed steganography method, and it was observed that the detector failed to detect the steg hidden inside those stego images.

**5.1.2 StegExpose Tool**

It is a statistical detector implemented on JAVA 1.6 and utilizes a detection algorithm in standard as well as fast fusion mode [32]. The stego image generated by the proposed algorithm is tested in fast fusion mode which is preferred most of the



**Fig. 4** Variation in payload capacity for 1-LSB (left) and 4-LSB (right) methods for Lung 5 as secret image

time for practical implementations. Fast fusion classification contains four detector stages, and in each stage for a newly added detector, the threshold value is analysed. The four detectors used are primary sets detector, Chi-square detector, sample pairs detector and RS analysis detector. If the mean value of all recently added detectors is above threshold, then the next stage detector is added and if the mean is below the threshold, then the image is classified as a clean image. In this work, a total 1000 images were combined in a 'testFolder' consisting of clean and stego images. The threshold value is chosen as 0.3, and it was observed that the other stego images with hidden content are detected except the 24 stego images which are generated using the proposed algorithm. Based on these results, it can be stated that the proposed algorithm is immune to Chi-square attack, RS attack and primary sets and sample pairs attack.

### 5.1.3 Histogram Analysis

The histogram of secret image Lung 5 before and after encryption is depicted in Fig. 6a, b, respectively. The uniform distribution of pixel intensities in the encrypted image affirms its robustness against histogram attack. The first-order statistics of the cover image is shown in Fig. 6c, and the resultant stego image histograms for 1-LSB, 2-LSB and 4-LSB methods are as shown in Fig. 6d–f, respectively. As per the previous work done by Juaid and Gutub [9], the 1-LSB implementation has resulted in more security in terms of less deteriorated histogram but reduced embedding strength. Whereas 4-LSB implementation is offering increased embedding strength at the cost of reduced security (high distortion in histogram). Thus, it can be suggested that a suitable trade-off between embedding strength and security should be made to achieve desired results.

Table 1 Visual imperceptibility metrics for various secret chest X-ray images

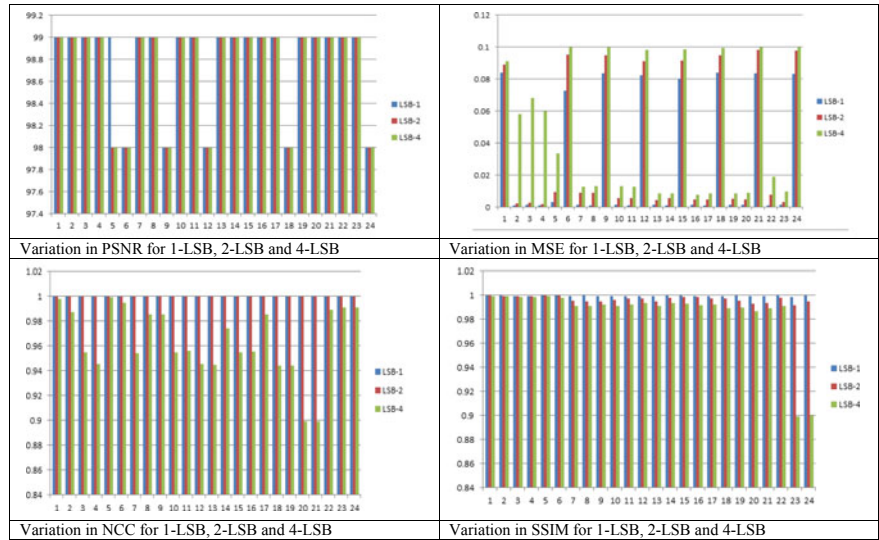
Fractal Cover used	Secret medical image used	For Stego image	PSNR			MSE			NCC			SSIM		
			1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB
Fractal 1.jpg with skewness = 0	Lung 1.jpg	1	99	99	99	0.0842	0.0889	0.0912	1.00	1.00	0.9977	0.9999	0.9997	0.9994
	Lung 2.jpg	2	99	99	99	0.0011	0.0024	0.0581	1.00	1.00	0.9874	0.9997	0.9993	0.9991
	Lung 3.jpg	3	99	99	99	0.0014	0.0026	0.0684	1.00	1.00	0.9547	0.9995	0.9991	0.9989
	Lung 4.jpg	4	99	99	99	0.0012	0.0021	0.0599	1.00	1.00	0.9457	0.9993	0.9991	0.9986
	Lung 5.jpg	5	99	98	98	0.0032	0.0093	0.0337	1.00	1.00	0.9995	0.9999	0.9996	0.9994
	Lung 6.jpg	6	98	98	98	0.0728	0.0952	0.0998	1.00	1.00	0.9947	0.9997	0.9997	0.9981
Fractal 2.jpg with skewness = 0	Lung 1.jpg	7	99	99	99	0.0014	0.0089	0.0128	1.00	1.00	0.9544	0.9992	0.9956	0.9911
	Lung 2.jpg	8	99	99	99	0.0011	0.0088	0.0133	1.00	1.00	0.9852	0.9998	0.9951	0.9910
	Lung 3.jpg	9	98	98	98	0.0838	0.0947	0.0997	1.00	1.00	0.9854	0.9993	0.9947	0.9921
	Lung 4.jpg	10	99	99	99	0.0014	0.0056	0.0132	1.00	1.00	0.9548	0.9995	0.9964	0.9914
	Lung 5.jpg	11	99	99	99	0.0011	0.0058	0.0129	1.00	1.00	0.9558	0.9995	0.9971	0.9922
	Lung 6.jpg	12	98	98	98	0.0824	0.0911	0.0981	1.00	1.00	0.9457	0.9993	0.9975	0.9934
Fractal 3.jpg with skewness = 0	Lung 1.jpg	13	99	99	99	0.0014	0.0044	0.0087	1.00	1.00	0.9447	0.9995	0.9946	0.9912
	Lung 2.jpg	14	99	99	99	0.0012	0.0057	0.0086	1.00	1.00	0.9741	0.9997	0.9978	0.9935
	Lung 3.jpg	15	99	99	99	0.0802	0.0915	0.0987	1.00	1.00	0.9547	0.9997	0.9988	0.9928
	Lung 4.jpg	16	99	99	99	0.0013	0.0047	0.0079	1.00	1.00	0.9554	0.9995	0.9987	0.9917
	Lung 5.jpg	17	99	99	99	0.0011	0.0049	0.0087	1.00	1.00	0.9854	0.9995	0.9974	0.9921
	Lung 6.jpg	18	98	98	98	0.0842	0.0947	0.0994	1.00	1.00	0.9445	0.9993	0.9971	0.9891

(continued)

Table 1 (continued)

Fractal Cover used	Secret medical image used	For Stego image	PSNR			MSE			NCC			SSIM		
			1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB	1-LSB	2-LSB	4-LSB
Fractal 4.jpg with skewness = 0	Lung 1.jpg	19	99	99	99	0.0014	0.0051	0.0087	1.00	1.00	0.9445	0.9996	0.9954	0.9899
	Lung 2.jpg	20	99	99	99	0.0013	0.0049	0.0091	1.00	1.00	0.8991	0.9993	0.9929	0.9868
	Lung 3.jpg	21	99	99	99	0.0838	0.0984	0.0997	1.00	1.00	0.8989	0.9993	0.9937	0.9891
	Lung 4.jpg	22	99	99	99	0.0012	0.0079	0.0189	1.00	1.00	0.9891	0.9997	0.9979	0.9911
	Lung 5.jpg	23	99	99	99	0.0013	0.0031	0.0097	1.00	1.00	0.9911	0.9988	0.9917	0.8994
	Lung 6.jpg	24	98	98	98	0.0833	0.0978	0.0997	1.00	1.00	0.9914	0.9997	0.9947	0.8995





**Fig. 5** Variation in various performance metrics for 1-LSB, 2-LSB and 4-LSB

**5.1.4 ML-Based Steganalysis**

As described in Sect. 3, a deep learning model Xception is trained with the features extracted from cover and stego images and are fed to WEKA tool for ML-based classification. Table 2 lists the prediction results of various ML classifiers. In these results, TP can be interpreted as true positive (model succeeds to predict the actual stego sample), TN can be interpreted as true negative (sample image is cover image and model does not predict as stego), FN can be interpreted as false negative (actual stego sample is not predicted as stego) and FP as false positive (actual cover sample is predicted as stego). We performed ten repetitions on ten-fold cross-validation on the cover and stego images. It can be observed that the J48 classifier gave the highest accuracy (81%) of true prediction among the other available classifiers. Thus, it can be interpreted that the trained model is offering sufficiently acceptable accuracy. This model is further used to detect the test images consisting 24 stego images created by proposed algorithm. It was noticed that the detector failed to steganalyze the hidden content inside those 24 stego images thus demonstrating the capability of proposed method to withstand the ML-based attacks.

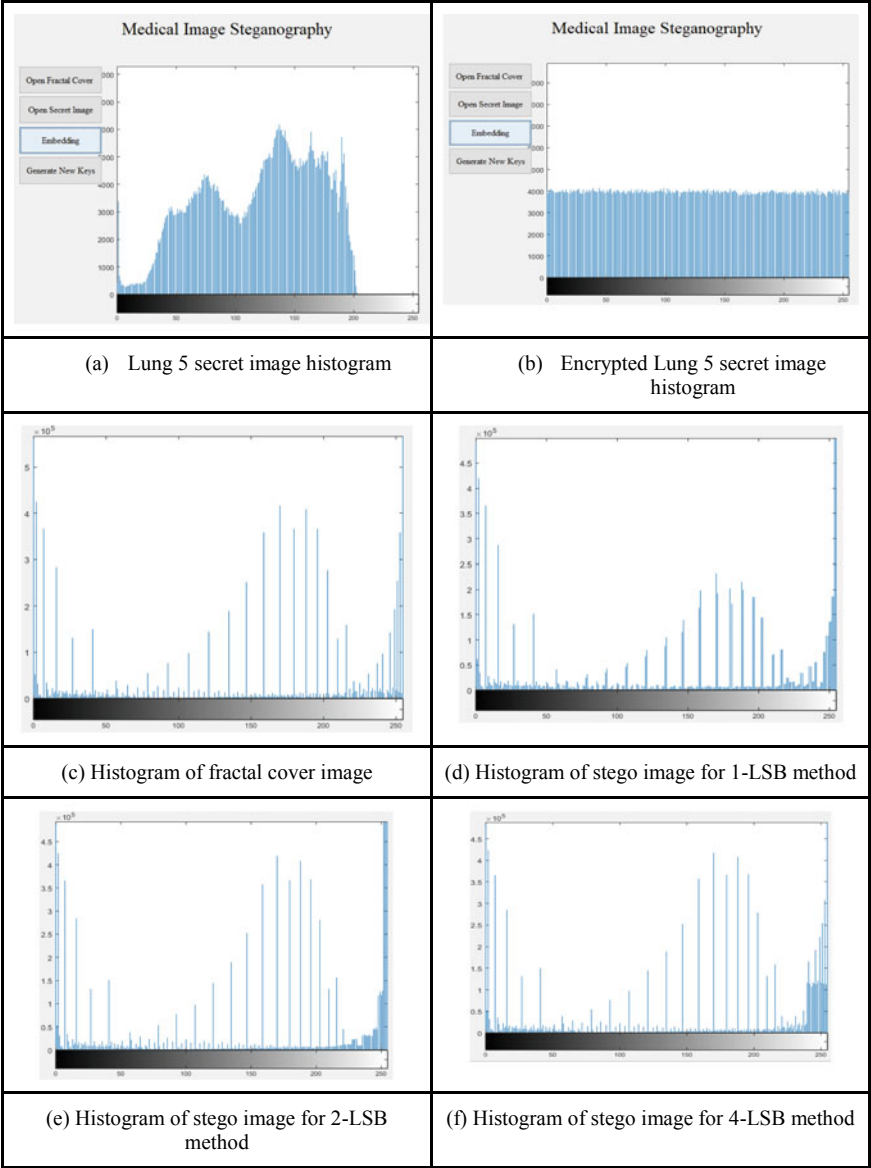


Fig. 6 Histogram analysis

**Table 2** ML classifiers prediction results

Classifier	TP rate	FP rate	% ROC value	% Accuracy
J48	0.81	0.18	84.1	<b>81.1</b>
Random Forest	0.72	0.27	78.2	72.5
Bayes Net	0.68	0.31	74.5	68.3

## 5.2 Comparison with Previous Work

The proposed algorithm is tested against SegSecret and StegExpose statistical steganalysis attacks and is found to perform better to resist them. Also, using ML-based steganalysis, promising results are obtained. The performance metrics are compared with the previous implementations of steganography and are depicted in Table 3. It shows that many steganographers have not paid enough attention towards cover selection criteria and its significance over the hardness of an image hiding algorithm. It can be seen that the proposed algorithm stands superior to the other methods in terms of security and imperceptibility.

## 6 Conclusion

A novel technique for hiding COVID-19 patients' secret medical images like CT scan, X-ray and other imaging modalities is developed using 1-LSB, 2-LSB and 4-LSB embedding methods with formula-driven fractal cover images. To enhance the security of secret data transmission, the modalities are first encrypted using RSA encryption and then concealed inside the unique covers. An extensive comparative evaluation is presented to demonstrate three different embedding methods of LSB insertion, and it can be inferred that a trade-off between security and visual perception can be achieved to build an optimum steganographic technique. Furthermore, the steg hiding capability of the proposed method is assessed by statistical as well as artificial intelligence-based steganalysis tools, and the obtained results validate the resistance competency against the possible security threats. In future, we plan to study several other steganalysis methods to develop a general detection method that can be used against any type of steganographic embedding process.

**Table 3** Comparison with previous work

Image used	Method used	PSNR	MSE	NCC	SSIM
Lena image	[1] LSB + RSA	42.88	0.05	–	–
Lena image	[2] LSB + RSA	63.38	0.0098	–	–
Lena image	[33] 3-LSB	35.70	–	–	–
Lena image	[34] LSB + chaotic encryption with Logistic map	55.98	0.1638	–	0.9986
Lena image	[35] Discrete Cosine Transform(DCT) + Elliptic Curve Cryptography(ECC)	42.12	–	–	0.9793
Fractal images	<b>Proposed method</b> 1-LSB, 2-LSB and 4-LSB + RSA + Fractal cover	99 (average value of value of 1-LSB, 2-LSB and 4-LSB methods)	0.0032 (average value 1-LSB, 2-LSB and 4-LSB methods)	1 (average value of value of 1-LSB, 2-LSB and 4-LSB methods)	0.9995 (average value of 1-LSB, 2-LSB and 4-LSB methods)

**Acknowledgements** One of us (VP) acknowledges the MATRICS Grant (MTR/2018/000203) from SERB, DST, Government of India.

## References

1. Kalaichelvi V, Meenakshi P, Vimala Devi P, Manikandan H, Venkateswari P, Swaminathan S (2020) A stable image steganography: a novel approach based on modified RSA algorithm and 2–4 least significant bit (LSB) technique. *J Ambient Intelligence Humanized Comput*. <https://doi.org/10.1007/s12652-020-02398-w>
2. Apau R, Adomako C (2017) Design of image steganography based on RSA algorithm and LSB insertion for android smartphones. *Int J Comput Appl* (0975–8887) 164(1)
3. Patidar V, Purohit G, Pareek NK (2018) A novel quasigroup substitution scheme for chaos based image encryption. *J Appl Nonlinear Dyn* 7(4):393–412. [\(2018\)](https://doi.org/10.5890/JAND.2018.12.007)
4. Teh JS, Alawida M, Sii YC (2020) Implementation and practical problems of chaos-based cryptography revisited. *J Information Security Appl* 50:102421. <https://doi.org/10.1016/j.jisa.2019.102421>
5. The JS, Alawida M, Sii YC (2020) Implementation and practical problems of chaos-based cryptography revisited. *J Information Security Appl* 50, 102421, ISSN 2214–2126. [\(2020\)](https://doi.org/10.1016/j.jisa.2019.102421)
6. Andreeva E, Mennink B, Preneel B (2015) Open problems in hash function security. *Des Codes Cryptogr* 77:611–631 [\(2015\)](https://doi.org/10.1007/s10623-015-0096-0)
7. Rout H, Mishra BK (2014) Pros and cons of cryptography, steganography and perturbation techniques. *IOSR J Electronics Commun Eng* (IOSR-JECE) e-ISSN: 2278-2834, p- ISSN: 2278–8735, pp 76–81
8. Aqeel I, Suleman MB (2019) A Survey on digital image steganography approaches. In: Bajwa I, Kamareddine F, Costa A (eds) *Intelligent technologies and applications. INTAP 2018. Communications in Computer and Information Science*, vol 932. Springer, Singapore. [https://doi.org/10.1007/978-981-13-6052-7\\_66](https://doi.org/10.1007/978-981-13-6052-7_66)
9. Al-Juaid N, Gutub A (2019) Combining RSA and audio steganography on personal computers for enhancing security. *SN Appl Sci* 1:830. <https://doi.org/10.1007/s42452-019-0875-8>
10. Reyad O, Karar ME (2021) Secure CT-image encryption for COVID-19 infections using HBBS-based multiple key-streams. *Arab J Sci Eng* 46 3581–3593. [\(2021\)](https://doi.org/10.1007/s13369-020-05196-w)
11. Kwee TC, Kwee RM (2020) Chest CT in COVID-19: what the radiologist needs to know. *RSNA Radiographics* 40(7). [\(2020\)](https://doi.org/10.1148/rg.2020200159)
12. Rousan LA, Elobeid E, Karrar M et al (2020) Chest x-ray findings and temporal lung changes in patients with COVID-19 pneumonia. *BMC Pulm Med* 20:245. <https://doi.org/10.1186/s12890-020-01286-5>
13. Yedrouj M (2019) *Steganalysis and steganography by deep learning*. Autre [cs.OH]. Université Montpellier, Français. ffnNT: 2019MONT095ff. ffile-02881987
14. Bose B, Dey D, Sengupta A, Mulchandani N, Patra A (2021) A novel medical image encryption using cyclic coding in Covid-19 pandemic situation. *J Phys: Conf Ser* 1797:012035. [\(2021\)](https://doi.org/10.1088/1742-6596/1797/1/012035)
15. Sahoo S, Sahoo SS (2020) A new COVID-19 medical image steganography based on dual encrypted data insertion into minimum mean intensity window of LSB of X-ray scans. In: 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, pp 1–6. [\(2020\)](https://doi.org/10.1109/INDICON49873.2020.9342067)
16. Prema T, Akkasaligar, Biradar S (2020) Selective medical image encryption using DNA cryptography. *Information Security J: Glob Perspect* 29(2):91–101. [\(2020\)](https://doi.org/10.1080/19393555.2020.1718248)

17. Paul AJ (2020) Recent advances in selective image encryption and its indispensability due to COVID-19. In: IEEE Recent Advances in Intelligent Computational Systems (RAICS), pp 201–207. <https://doi.org/10.1109/RAICS51191.2020.9332513>. (2020)
18. Zou Y, Zhang G, Liu L (2019) Research on image steganography analysis based on deep learning. *J Vis Commun Image R* R60:266–275. <https://doi.org/10.1016/j.jvcir.2019.02.034>. (2019)
19. Jung K-H (2019) A study on machine learning for steganalysis. In: Proceedings of the 3rd international conference on machine learning and soft computing—ICMLSC 2019. <https://doi.org/10.1145/3310986.3311000>. (2019)
20. Pichardo-Méndez JL, Palacios-Luengas L, Martínez-González RF et al (2020) LSB pseudo-random algorithm for image steganography using skew tent map. *Arab J Sci Eng* 45:3055–3074. <https://doi.org/10.1007/s13369-019-04272-0>
21. Parah SA, Ahad F, Sheikh JA, Bhat GM (2017) Hiding clinical information in medical images: a new high capacity and reversible data hiding technique. *J Biomed Informatics* 66:214–230. ISSN 1532-0464. <https://doi.org/10.1016/j.jbi.2017.01.006>
22. Wardlaw WP (2000) The RSA public key cryptosystem. In: Coding theory and cryptography, pp 101–123. [https://doi.org/10.1007/978-3-642-59663-6\\_6](https://doi.org/10.1007/978-3-642-59663-6_6)
23. Asha D, Vinod P (2020) Development and analysis of IWT-SVD and DWT-SVD steganography using fractal cover. *J King Saud Univ—Comput Information Sci*. <https://doi.org/10.1016/j.jksuci.2020.10.008>
24. [www.ultrafractal.com](http://www.ultrafractal.com)
25. Molato MRD, Gerardo BD (2018) Cover image selection technique for secured LSB-based image steganography. In: Proceedings of the 2018 international conference on algorithms, computing and artificial intelligence—ACAI 2018. <https://doi.org/10.1145/3302425.3302456>
26. Dalal M, Juneja M (2020) Steganography and Steganalysis (in digital forensics): a cybersecurity guide. *Multimedia Tools Appl* 80:5723–5771. <https://doi.org/10.1007/s11042-020-09929-9>
27. Amirtharajan R, Adharsh D, Vignesh V, John Bosco Balaguru R (2010) PVD blend with pixel indicator —OPAP composite for high fidelity steganography. *Int J Comput Appl* (0975–8887) 7(9)
28. <https://hiroksukataoka16.github.io/Pretraining-without-Natural-Images/#dataset>
29. [www.Mathworks.com](http://www.Mathworks.com)
30. MATLAB (2005) The language of technical computing. The Mathworks, Programming version 7
31. <https://github.com/ieee8023/covid-chestxray-dataset/tree/master/images>
32. Benedikt Boehm (2014) StegExpose—a tool for detecting LSB steganography
33. Jung K-H (2019) A study on machine learning for steganalysis. In: ICMLSC 2019: Proceedings of the 3rd international conference on machine learning and soft computing, pp 12–15. <https://doi.org/10.1145/3310986.3311000>
34. Kasapbasi MC (2019) A new chaotic image steganography technique based on Huffman compression of Turkish texts and fractal encryption with post-quantum security. *IEEE Access* 7:148495–148510
35. Duan X, Guo D, Liu N, Li B, Gou M, Qin C (2020) new high capacity image steganography method combined with image ECC and Deep Neural Network. *IEEE Access* 8:25777–25788. <https://doi.org/10.1109/ACCESS.2020.2971528>

# Text Summarization of News Articles



Tanvi Oka, Prachi Patankar, Shivani Rege, and Mrudul Dixit

**Abstract** Text summarization gives condensed information of long texts and documents. Extractive and abstractive are two approaches for summarizing texts. Extractive text summarization is a type of forming summary by finding out the key phrases in a text or article. These important phrases are extracted to form the summary. In this paper, we aim to carry out text summarization of texts in two foreign languages, namely English and German, and two Indian languages: Marathi and Hindi. We have considered an entire sentence in a text as a feature to be extracted to form the summary. We have summarized news articles in the above-mentioned languages. For summarization, supervised or unsupervised algorithms can be used. We aim to find performance of unsupervised algorithms, namely weighted, TextRank, and fuzzy logic for summarizing texts in the above-mentioned languages. The performance evaluation of algorithms is done using ROUGE metric.

**Keywords** Text summarization · Weighted algorithm · TextRank algorithm · Fuzzy logic algorithm · Natural language processing · Indian languages · ROUGE metric

## 1 Introduction

The Internet consists of web pages, news articles, status updates, blogs, and so much more. People enjoy quick access to this enormous amount of information. However, most of this information is redundant, insignificant, and may not convey the intended

---

T. Oka (✉) · P. Patankar · S. Rege · M. Dixit  
MKSSS'S Cummins College of Engineering for Women, Pune, India  
e-mail: [tanvi.oka@cumminscollege.in](mailto:tanvi.oka@cumminscollege.in)

P. Patankar  
e-mail: [prachi.patankar@cumminscollege.in](mailto:prachi.patankar@cumminscollege.in)

S. Rege  
e-mail: [shivani.rege@cumminscollege.in](mailto:shivani.rege@cumminscollege.in)

M. Dixit  
e-mail: [mrudul.dixit@cumminscollege.in](mailto:mrudul.dixit@cumminscollege.in)

meaning. Here, text summarization comes into picture which is capable of extracting useful information and leaves out inessential and insignificant data. Implementing summarization enhances the readability of documents, reduces the time spent in researching the information, and allows more information to be fitted in a particular area. Text summarization is a challenging field of natural language processing (NLP) as it requires precise text analysis which includes semantic and lexical analysis to form an accurate summary.

The database used is BBC news dataset for the English language. This dataset is already divided into five categories. There are a total of 4450 records. One of the advantages of this database is that it comes with a set of summaries as well. This is helpful to check the accuracy of the results at the end. For Marathi language, we have taken a dataset from GitHub. This dataset also comes with a set of summaries. For Hindi and German language news, we have created our own dataset by taking news from different news Web sites.

The language used is Python. Since text summarization is a subdomain of natural language processing, Python's natural language toolkit is needed. It helps to work with human language data to build Python programs. It contains packages to convert human language into machine language and give appropriate responses to it. NLTK supports classification, tokenization, stemming, tagging, parsing, and semantic reasoning.

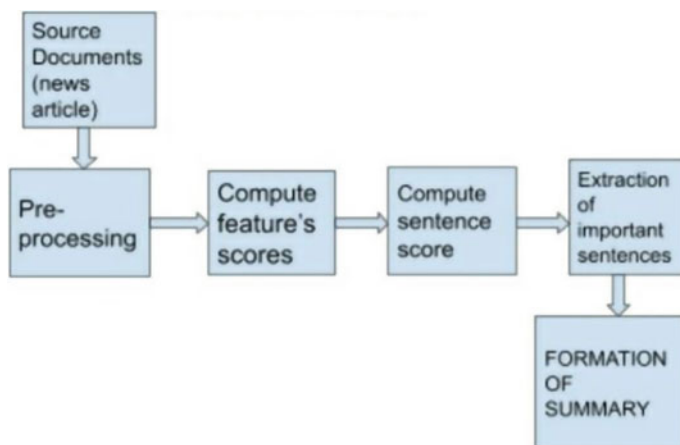
Text summarization uses two approaches, namely extractive and abstractive [1]. In abstractive summarization technique, we extract the main points from the text and generate new sentences to give the summary. While in extractive summarization, we extract important sentences by keeping the facts intact. Here, we do not generate new sentences. Instead, we just extract the important sentences as it is. Hence, we have used an extractive summarization approach.

Along with English language, we are going to summarize news articles in regional languages such as Hindi and Marathi [2]. We are also going to summarize news articles in German language [3]. For languages other than English, we need a Python library-codecs. Python comes with a number of codecs built-in, either implemented as C functions or with dictionaries as mapping tables. Many of the character sets support the same languages. They vary in individual characters and in the assignment of characters to code positions.

## 2 Block Diagram

See Fig. 1.





**Fig. 1** Block diagram of system

### 3 Algorithm

In this system, text summarization is implemented using three different algorithms—weighted, TextRank, and fuzzy. Brief information regarding these algorithms is given below:

#### 3.1 *Weighted Algorithm*

Weighted algorithm is an unsupervised algorithm [4]. This algorithm helps us to find important words in a text document which helps determine important sentences.

The steps for finding the scores are as follows:

First, sentence tokenization is done. Then, the title of the news article is separated. Case of the text is lowered. Next, word tokenization is used on these sentences, and the stopwords are removed. After the stopwords are removed, next is computation of the frequency for each distinct word occurring in the text. The highest frequency in the list of frequencies mentioned above is found out. To find the score of each word, the frequency of that word is divided by the highest frequency. After finding the score of each distinct word, the total score of each sentence is found out by adding the scores of each word in that sentence. After finding the score of each sentence, extract the highest scoring sentences to form the summary. While doing this, it is assumed that the summary is 25% of the total text.

### 3.2 *TextRank Algorithm*

It is a graph-based ranking model for text processing [5]. It helps to find the keywords and also the relevant sentences in a text document. TextRank is an unsupervised extraction algorithm.

The steps for finding the scores are as follows:

First, split the sentences. The next step is sentence tokenization. After tokenization of sentences, preprocessing is done. Here, the first step is lowering the case of all the sentences. Once all the sentences are in lower case, we remove the numerical data. Next, the stopwords are removed, and sentence vectorization is done. After this, find similarity between the sentences. Then, convert the similarity matrix into a graph, and rank the sentences [6]. Once ranking of the sentences is done, select the top sentences which will form the summary of the news article.

### 3.3 *Fuzzy Algorithm*

Fuzzy logic is a method of reasoning that resembles human reasoning [7]. This approach imitates the way of decision making in humans that involves all intermediate possibilities between digital values YES and NO.

The first step is splitting the text into sentences. After splitting, preprocessing is done which includes sentence tokenization, lowering the case of each sentence, and removing the stopwords. For this algorithm, consider the following eight features:

Feature 1: Title feature—This feature would be considered only if the given document/text has a title. If the given document does not have a title, we directly go to feature 2. This feature is used to find similarity between title sentence and every other sentence in the document. This is determined by counting the number of common content words—sentence and the title.

The short sentences such as author names and date lines are not expected to belong to summary. So, this feature is to filter such sentences.

Feature 2: Sentence length—Here, the length of each sentence is found. To find the feature score of each sentence, divide the length of each sentence with the length of the maximum sentence.

Feature 3: Sentence position—Score of the sentence will depend on the position of the sentence in the text, i.e., sentence occurring first will have a higher rank than the last sentence. The score of each sentence is the rank of the sentence divided by the number of sentences in the document.

Feature 4: Numerical data—Sentences containing numerical data contain important information/statistics related to the document/text. To find the score of this feature for each sentence, divide the number of numerical data in it by its length.

Feature 5: Thematic words—This feature helps to understand the theme of the document by considering the most occurring words. Here, first find the frequency of each unique word in the document. Then, arrange them in descending order of

frequency. Then, consider the first five words. The score of each sentence will be counted by adding the occurrence of the first five words in each sentence.

**Feature 6: Sentence to sentence similarity**—This feature is used to find the similarity between a sentence and every other sentence. In this,  $n$ -by- $n$  matrix is formed, where  $n$  is the total number of sentences in a document. As the sentence should not be compared to itself, set the diagonal elements to zero. This matrix is then converted to a graph, and the sentences are ranked.

**Feature 7: Term weight**—Compute the frequency for each distinct word occurring in the text, and find the highest frequency amongst them. To find the score, add the frequency of each word in that sentence, and divide it by the maximum summation of all sentences in the document.

**Feature 8: Proper nouns**—Having proper noun in a sentence enhances the meaning, and hence, having a proper noun in a sentence is considered important. With the help of POS tagging, we identify the proper nouns in a sentence. Then, compute the score by adding the total number of proper nouns in a sentence divided by the sentence length.

After getting eight features, a fuzzy control system is created, which models how you select important sentences for the summary. While finding the important sentences, the feature scores are between 0 and 10, the importance value will also be between 0 and 10. Antecedent (input) objects are the features. While the consequent (output) object is the importance of sentence, then membership functions are built to graphically represent the fuzzy set. Based on feature values obtained for each sentence, the fuzzy rules are set. Once the rules are set, we find the importance of each sentence. Then, the sentences with higher importance are selected as summary.

## 4 Result

In this paper, we implement content evaluation to evaluate our system-generated summary [8]. ROUGE stands for recall-oriented understudy for gisting evaluation. It tells how much similar is the system-generated summary to the reference (ideal) summary created by humans [9]. Following tables show average F1 score, precision, and recall values of the three algorithms calculated for 25 news articles for each of the four selected languages.

F1 score is the harmonic mean of precision ( $P$ ) and recall ( $R$ ) (Tables 1, 2, 3, and 4; Figs. 2, 3, 4, and 5).

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \quad (1)$$

Graphical Representation of the above tables is as follows:

**Table 1** F1 score, precision, and recall values of three algorithms applied on English news dataset

		Weighted	TextRank	Fuzzy Logic
R1	F1s	0.63	0.64	0.54
	P	0.78	0.93	0.83
	R	0.56	0.54	0.49
R2	F1s	0.55	0.63	0.54
	P	0.68	0.93	0.74
	R	0.47	0.46	0.43
RL	F1s	0.65	0.69	0.64
	P	0.75	0.96	0.80
	R	0.57	0.55	0.54

**Table 2** F1 score, precision, and recall values of three algorithms applied on German news dataset

		Weighted	TextRank	Fuzzy Logic
R1	F1s	0.52	0.46	0.52
	P	0.54	0.62	0.60
	R	0.53	0.39	0.47
R2	F1s	0.41	0.35	0.41
	P	0.43	0.47	0.49
	R	0.42	0.30	0.38
RL	F1s	0.50	0.45	0.51
	P	0.52	0.58	0.57
	R	0.51	0.38	0.47

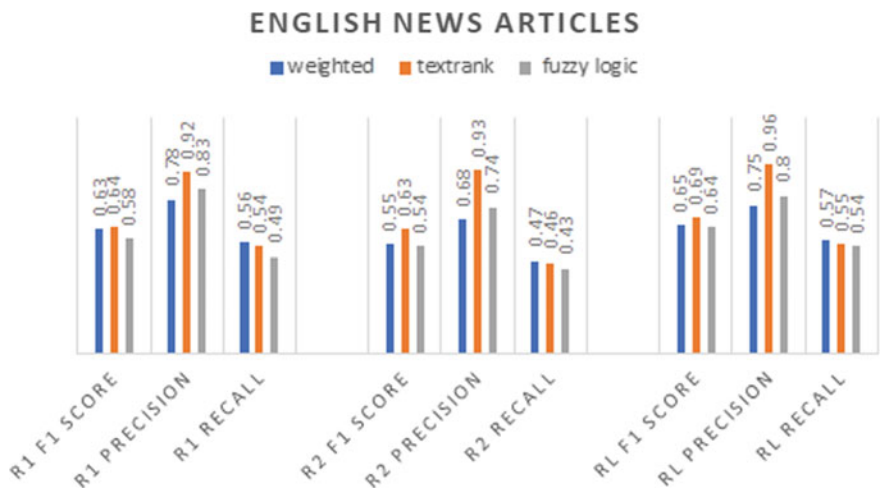
**Table 3** F1 score, precision, and recall values of three algorithms applied on Marathi news dataset

		Weighted	TextRank	Fuzzy Logic
R1	F1s	0.53	0.52	0.43
	P	0.66	0.61	0.65
	R	0.52	0.50	0.34
R2	F1s	0.48	0.46	0.36
	P	0.59	0.51	0.57
	R	0.46	0.44	0.32
RL	F1s	0.55	0.54	0.43
	P	0.67	0.63	0.65
	R	0.52	0.51	0.36

**Table 4** Shows F1 score, precision, and recall values of three algorithms applied on Hindi news dataset

		Weighted	TextRank	Fuzzy Logic
R1	F1s	0.76	0.57	0.69
	P	0.73	0.60	0.69
	R	0.79	0.55	0.70
R2	F1s	0.69	0.46	0.61
	P	0.67	0.48	0.61
	R	0.72	0.45	0.61
RL	F1s	0.715	0.51	0.67
	P	0.74	0.54	0.68
	R	0.77	0.48	0.67

<sup>a</sup> R1 = Rouge 1 R2 = rouge 2 RL = Rouge-L, F1s = F1 score P = Precision R = Recall



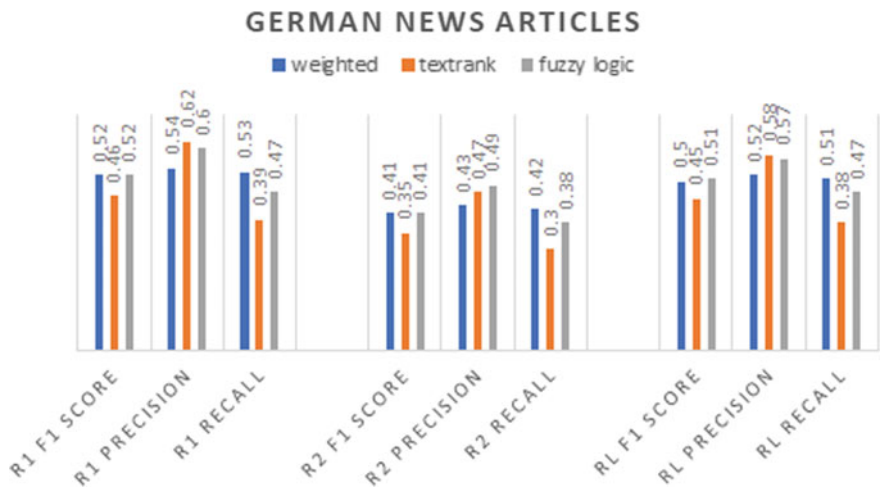
**Fig. 2** Comparative analysis of weighted, TextRank, and fuzzy logic algorithm for English news

## 5 Conclusion

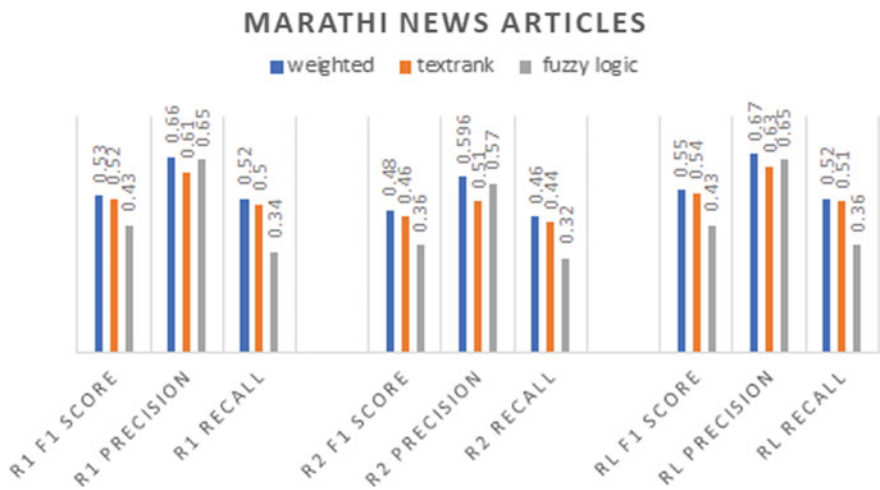
This paper focuses on performance of unsupervised algorithms which are weighted frequency, TextRank, and fuzzy logic for summarizing news articles.

For English news, the TextRank algorithm shows better performance, with average F1 score for R1, R2 and RL more than average F1 scores of R1, R2, and RL for weighted and fuzzy logic.

For German news, weighted and fuzzy logic show similar F1 scores of R1, R2, and RL, which are comparatively more than F1 scores of R1, R2, and RL for TextRank algorithm.



**Fig. 3** Comparative analysis of weighted, TextRank, and fuzzy logic algorithm for German news

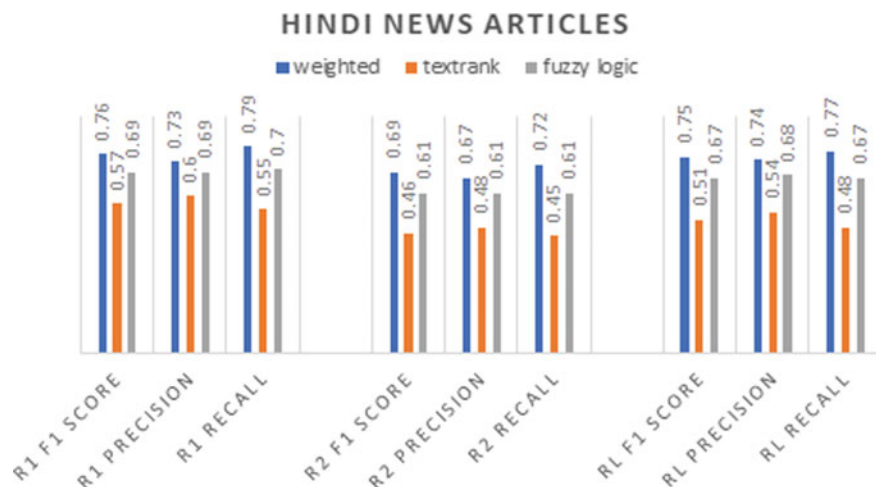


**Fig. 4** Comparative analysis of weighted, TextRank, and fuzzy logic algorithm for Marathi news

For Marathi news, weighted algorithm shows better performance, with average F1 score for R1, R2, and RL more than average F1 scores of R1, R2, and RL for TextRank and fuzzy logic algorithms.

For Hindi news, weighted algorithm shows better performance, with average F1 score for R1, R2, and RL more than average F1 scores of R1, R2, and RL for TextRank and fuzzy logic algorithms.

There are also supervised approaches for text summarization. They can be applied for better performance for Indian languages [10].



**Fig. 5** Comparative analysis of weighted, TextRank, and fuzzy logic algorithm for Hindi news

## 6 Future Scope

Text summarization is an application which can be used in various fields and will gain more and more importance with the growing amount of data each day available online.

Web scraping can be used for directly accessing news articles available on the web pages.

Multi-document text summarization can also be done for summarizing news articles of a specific category, for e.g., business, politics, sports, technology, entertainment, environment, etc., rather than summarizing a single news article each time from that category.

## References

1. Moratanch N, Chitrakala S (2017) A survey on extractive text summarization, ICCSP
2. Bhosale S, Joshi D, Bhise V, Deshmukh RA (2018) Marathi e-newspaper text summarization using automatic keyword extraction technique. *Int J Adv Eng Res Dev* 5(03)
3. Sherry, Saini A (2017) Multilingual text summarization techniques. *IJRET* 06(06)
4. Sri Rama Raju Bhargav Allarpu T (2017) Text summarization using sentence scoring method. *IRJET* 4(4)
5. Mihalcea R, Tarau P. TextRank: bringing order into texts. Department of Computer Science University of North Texas
6. Tanwi, Ghosh S, Kumar V, Jain YS, Avinash B (2014) Automatic text summarization using Text Rank. *IRJET* 6(4)
7. Pallavi, Patil D, Kulkarni NJ (2014) Text summarization system using Fuzzy Logic. *IJIRAE* 1(3)

8. Steinberger J, Jeřek K (2009) Evaluation measure for Text Summarization. *Comput Informatics* 28
9. Lin C-Y (2004) ROUGE: a package for automatic evaluation of summaries. *ACL Anthology*, Vol: Text summarization branches out
10. Verma P, Verma A (2020) Accountability of NLP tools in text summarization for Indian languages. *J Sci Res* 64(1)



# Simple Square Slot Wearable Antenna with Coaxial Feed for ISM Band Applications at 5.3 GHz Frequency



Vandana Yadav, Akhilesh Kumar, Prakhar Yadav, Amrees Pandey, and R. S. Yadav

**Abstract** This paper presents a wearable microstrip patch antenna  $17 \times 34 \times 1.092$  mm<sup>3</sup> for industrial scientific and medical application (ISM) at 5.3 GHz frequency. This is a very simple design of a wearable antenna with high gain and improved return loss for this frequency. This antenna acts as communicating device to send or receive signals with sensors. Due to simple square slot, the required frequency 5.3 GHz is polarized. Different garments can be used as substrate for wearable antenna. Polyester is used as dielectric material in three layers of that to obtain better characteristics of the antenna. This antenna has better polarization with the coaxial feed at the desired frequency. Proposed antenna is simulated on ANSYS HFSS version 17. This wearable antenna structure resonates at 5.3 GHz with an improved reflection coefficient of  $-30$  dB, a maximum gain of 7.21 dBi, and bandwidth of 400 MHz. This antenna is applicable for wireless communication, ISM band, WLAN, and Wi-Fi.

**Keywords** Square slot · Microstrip · Wearable antenna · Coaxial feed · Polyester · Garments · ISM · WBAN

## 1 Introduction

In recent years, the development of wearable devices helps us to improve the grade of our lives by continuous monitoring of wearer through wireless communication. There are plenty of applications in wearable antenna especially, in the field of healthcare, security, sports, space, rescue, glamor, and wireless communication. In fifth generation, wearable antenna is an integral part of Internet of things (IoT) to connect various things such as wireless body area networks (WBAN) application are shown [1]. Due to its lightweight and flexibility, this wearable antenna can be easily worn by a person

---

V. Yadav (✉) · P. Yadav · A. Pandey · R. S. Yadav

Department of Electronics and Communication, University of Allahabad, Prayagraj, UP 211002, India

A. Kumar

Electronics Engineering Department, Institute of Engineering and Rural Technology, Prayagraj, UP 211004, India

and observer can communicate him with cellular network. So far many researchers have designed various wearable antennas for different applications [2] on different garments such as on kurta, shirts, and jackets. In [3], a slot antenna is made for wearable medical application, fed by coplanar wave guide to enable live monitoring of wearer’s health. Using metamaterial is another way to improve the gain and return loss [4]. The wearable antennas play important role as they are lightweight and flexible as required for GPS receivers and mobile phones applications [5]. Fabric selection for substrate is the major factor to design best textile antenna such as dielectric material, yarn density, and pattern of weaving. There are various fabrics which can be used as a substrate for example polyester, cotton, liquid crystal polymer, Nomex, copper coated fabric, Nylon, etc. [6]. Polyester is the best among all because of its low dielectric material and moisture gain value. The textile antenna should not have to effect on the health of the human body; so substrate materials are a very essential choice. From literature survey, it has been observed that thicker substrate improved antenna’s characteristics, and multilayering a substrate is one of the best ways to obtain adequate height of the antenna [8]. Thicker substrate minimizes the leakage current, increases the bandwidth, and reduces the Q-factor. The comparative analysis of proposed antenna with the reported antennas is show in Table 1. It can be clearly observed from Table 1 that the proposed antenna is equivalent to [7] whereas smaller in size than [8–11]. Gain of proposed antenna is better than [7, 8, 11] whereas is less than [9, 10]. As illustrated from Table 1, the proposed antenna is best among all reported antenna as it is smaller in size with improved gain, with bandwidth of 7.5%. The presenting antenna is simplest structure among all the reference antennas.

In this paper, three layers of polyester are added to obtain optimum height of the substrate. A square slot is cut on the patch to resonate the required frequency. In textile antenna bending effect is the demerits that can be also minimized with this slot. The feeding of this antenna is 50 Ω MSA coaxial fed. The coaxial feed is the

**Table 1** Comparative analysis of proposed antenna with reported antennas

Reference antenna	Size (mm)	Gain (dBi)	Operating frequency	Frequency range (GHz)	Bandwidth %	Special with
7	15 × 30	4.95	5.1	5–5.2	2	Embedded with metamaterial
8	20.5 × 94.5	6.28	5.3	3.35–3.45	2.9	E-shaped patch
9	45 × 52	8.6	2.45	2.43–2.46	1.2	Jeans as substrate
10	97 × 97	8	2.45	2.45–2.54	1	AMC reflector is used
11	39 × 46	5.04	2.5	2.38–2.5	3.6	Multilayer in substrate
Proposed antenna	15 × 30	7.21	5.3	5.1–5.5	7.5	Square cut

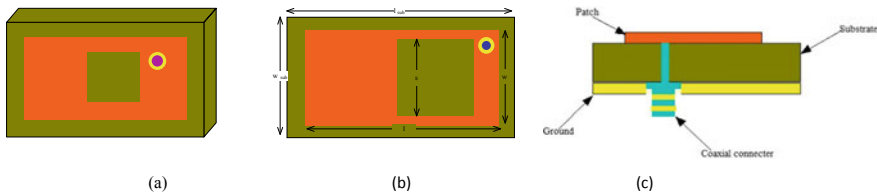
best among all feeding techniques to enhance the gain, return loss, and bandwidth [12]. The main advantage of this coax feed is that the feed location is placed at any point on the patch of the antenna to match the impedance of cable and antenna input. The purpose of this wearable antenna is to improve the gain, return loss, and bandwidth for the ISM band. The attractive features of this antenna are low cost, lightweight, portable, flexible, and low profile structure for wireless communication for example we can monitor the wearer by a base station in a particular area. The proposed antenna is polarized at desired frequency for wireless communication that is 5.3 GHz with a gain of 7.21 dBi, bandwidth 400 MHz, and returns loss  $-30$  dB.

## 2 Antenna Configurations and Analysis

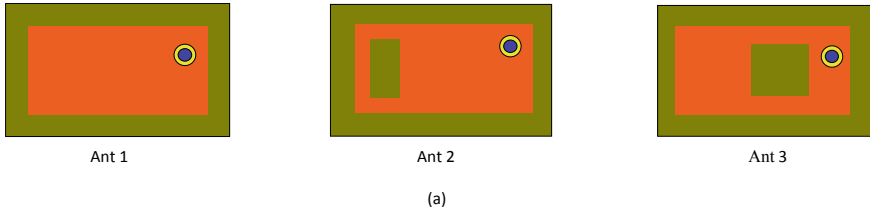
This square-shaped wearable antenna is designed with three layers: one is conducting rectangular patch on top, polyester substrate is in the middle, and conducting ground on the bottom of the structure [13]. In this paper, polyester is used as substrate material in layered form to increase the height of the substrate. The geometry of this wearable antenna is shown in Fig. 1. Microstrip rectangular patch antenna is a very simple structure in the designing of the microstrip patch antenna [14]. The calculation of all dimensions of the simple microstrip patch antenna is taken [15] from where we calculated the length and width of the patch using Eq. (1).

$$fo = \frac{c}{2\sqrt{\epsilon e}} \left[ \left( \frac{m}{l} \right)^2 + \left( \frac{n}{w} \right)^2 \right]^{1/2} \quad (1)$$

where  $fo$  is the resonant frequency,  $\epsilon e$  is effective dielectric constant,  $m$  and  $n$  are the modes as  $TM_{10}$ ,  $l$  is length and  $w$  is the width of the rectangular patch, and  $c$  is the light velocity in Eq. (1). The effective dielectric constant is calculated with the help of Eq. (2) where  $\epsilon r$  is the relative permittivity of the substrate ( $\epsilon r = 2.3$ ) for the polyester substrate.



**Fig. 1** **a** Structure of proposed wearable antenna, **b** layout of the antenna, and **c** side view of the wearable antenna with coaxial feed description



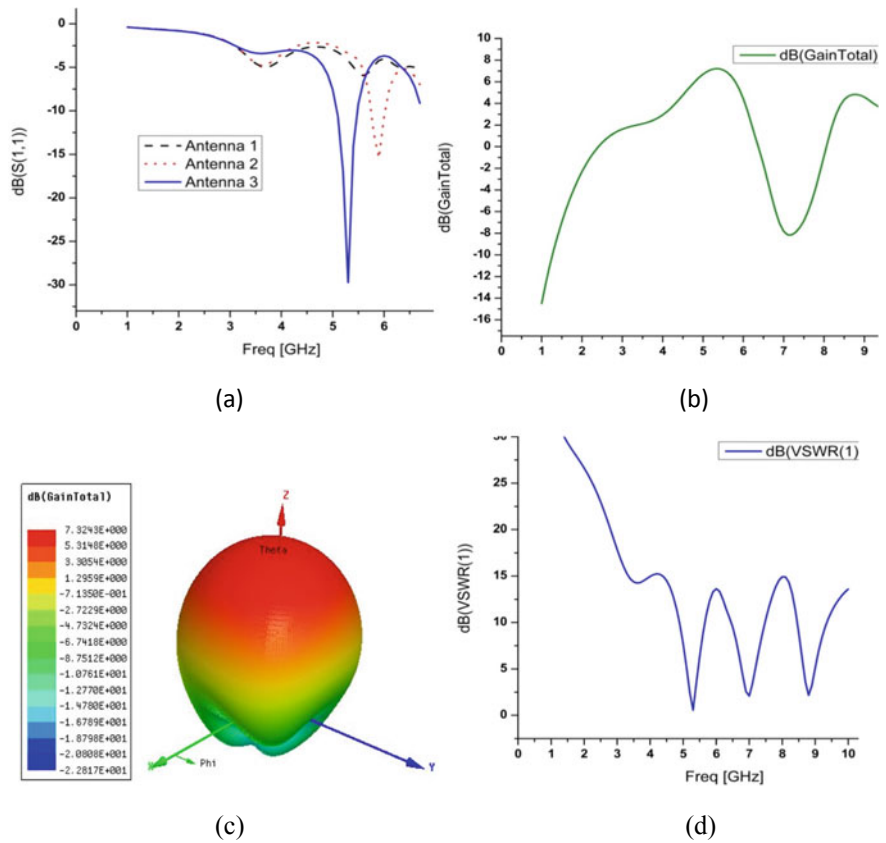
**Fig. 2 a** Evolution of design of wearable antenna as Ant1, Ant2, and Ant3

$$\varepsilon e = \frac{\varepsilon r + 1}{2} + \frac{\varepsilon r - 1}{2} \left[ 1 + 12 \frac{h}{w} \right]^{-1/2} \quad (2)$$

The length ( $l$ ) of the patch is 30 mm and the width ( $w$ ) is 15 mm. To obtain desired frequency 5.3 GHz, a square slot  $10 \times 10 \text{ mm}^2$  is cut in the rectangular patch. This slot also reduces the conducting material of the antenna. The thicker substrate is good for antenna design so to obtain an optimum height of the substrate, three layers of polyester are added for better performance of this wearable antenna. The dielectric constant of the polyester material is 2.3, and the height is 1.092 mm. The dimension of the substrate is  $17 \times 34 \text{ mm}^2$  and at the top of the structure is a conducting patch having geometry of  $15 \times 30 \text{ mm}^2$ . Conducting patch can be printed on textile by screen printing that will be a cost-effective option to create a patch on the substrate. Progress of the antenna is shown in Fig. 2, when the simple rectangular patch is designed and simulated, then this structure is polarized at 7.22 GHz with the gain of 0.44 dBi and the return loss is  $-17 \text{ dB}$ , that is shown in Fig. 2a as ant1. The required frequency could not obtain with ant1, then a rectangular slot is cut on the patch opposite corner of feed point to lower the polarized frequency this structure also cannot provide the desired result that is shown in Fig. 2b as Ant2. To get proper impedance matching for the WLAN frequency range, a square slot is created in the patch and this structure of antenna is shown as Ant3 in Fig. 2c that is proposed antenna. The overall dimensions of the proposed antenna are shown in Fig. 1b. The proposed antenna has a very simple structure with high gain and minimum return loss. The feed type and feed position also have importance for proper impedance matching in antenna design [16]. The coaxial feed is a  $50 \Omega$  SMA connector that should be at the  $L/2$  position along the length. For proper impedance matching, the actual feed point is at  $y = 11 \text{ mm}$  and  $x = 4 \text{ mm}$  in this antenna design. Proposed antenna is polarized at this frequency with improved gain and return loss due to coaxial feed in the design of antenna. The overall dimensions of this wearable antenna are shown in Table 2. And the step-by-step progress of this design is shown in Fig. 2a–c, as Ant1, Ant2, and proposed Ant, all result of these three antennas are also analyzed and compared results are shown in Fig. 3a, where we can see that square slot reduces the polarization frequency. Position of the square slot and coaxial feed point is the main factor to obtain the 5.3 GHz frequency with gain of 7.21 dBi and return loss  $-30 \text{ dB}$ . The square slot also reduces the patch dimension and avoids the bending effect due to body movement. Dielectric material polyester improves

**Table 2** Wearable microstrip patches antenna specifications

Layers design parameters dimensions (mm)		
Polyester substrate	Thickness ( $h$ )	1.092
	Length ( $l_{\text{sub}}$ )	34
	Width ( $w_{\text{sub}}$ )	17
	Dielectric constant ( $\epsilon_r$ )	2.3
	Dielectric loss ( $t_{\text{and}}$ )	$1.2 \times 10^{-2}$
Rectangular patch	Length ( $l$ )	30
	Width ( $w$ )	15
	Slot ( $s$ )	10
Ground	Length	34
	Width	17



**Fig. 3** **a** Return loss of wearable polyester ant1, ant2, and ant3, **b** simulated gains of the wearable polyester ant3, **c** simulated three-dimensional radiation pattern of ant3, and **d** VSWR of the proposed ant3

the antenna characteristics. This wearable antenna is simulated on ANSYS High Frequency Structure Simulator version 17 because this is better simulation software as compared to all electromagnetic wave simulators.

### 3 Simulation Results and Discussions

The simulated results of this wearable antenna are return loss, gain, radiation pattern, and VSWR which are drawn concerning the frequency. The bandwidth of this proposed antenna is 400 MHz considering under the return loss -10 dB. The return loss of the proposed antenna is shown in Fig. 3a where  $S_{11}$  goes to -30 dB at 5.3 GHz that is the best performance in the textile antenna with a simple structure. The  $s$  parameter of this antenna shows that minimum power is returned so its performance is better for the ISM band. Without a slot, this antenna cannot be polarized with this frequency range. In Fig. 3b, the proposed wearable antenna resonates with a 7.21 dBi gain at a 5.3 GHz frequency range. The simulated 3D radiation pattern of this antenna is shown in Fig. 3c. From this figure, we can see the maximum radiation is obtained at 5.3 GHz and the red portion of this polar diagram shows the maximum gain is obtained for this particular bandwidth that means maximum power is transferred in that range of frequency. It is clear that moisture absorption is the big problem with textile substrate antenna due to humidity in the air, skin sweat snow because water has more dielectric constant than fabric. This problem can be resolved by dip coating with PVB solution. The proposed layered polyester substrate antenna showed the best impedance matching with the desired frequency range. This proposed wearable antenna is polarized at 5.3 GHz frequency with the gain of 7.21 dBi and return loss -30 dB, and all results are analyzed and shown in Fig. 3.

### 4 Conclusion

The proposed wearable polyester base antenna is designed for 5.3 GHz with enhanced bandwidth of 400 MHz. To increase the height of the substrate, three layers of polyester are added; hence, a significant effect is obtained with this height of the substrate. The size and position of a slot on the patch help to polarize the required frequency with improved gain 7.21 dBi and  $S_{11}$  is -30 dB. All these results can be possible with three factors: one is the height of the substrate and second is the slot on the square patch and the third is feed position. There are various substrate materials but polyester is the best among all. This antenna can be used on the jacket or other clothes that are made of polyester fabrics. This wearable antenna is best suited for of ISM band applications. The main limitation of this design will be when we will go for fabrication due to polyester substrate material that is durability of the antenna. In the future, this structure can be updated with embedded metamaterial for various applications. The partial ground will be also an option to improve the gain

and another parameter of this wearable antenna. Different substrate materials can be used for various applications such as for security, medical, and communications.

## References

1. Turkmen M, Yalduz H (2018) Design and performance analysis of a flexible UWB wearable textile antenna on jeans substrate. *Int J Information Electronics Eng* 8(2):15–18
2. Mistry KK et al (2018) A design of elliptical edge-fed circularly polarized patch antenna for GPS and Iridium applications. In: 2018 2nd URSI Atlantic Radio Science Meeting (AT-RASC). IEEE
3. Li YJ, Lu ZY, Yang LS (2019) CPW-fed slot antenna for medical wearable applications. *IEEE Access* 7:42107–42112
4. Jiang, Hao Z et al (2014) A compact, low-profile metasurface-enabled antenna for wearable medical body-area network devices. *IEEE Trans Antennas Propag* 62(8):4021–4030
5. Paracha, Nisar K et al (2019) Wearable antennas: a review of materials, structures, and innovative features for autonomous communication and sensing. *IEEE Access* 7:56694–56712
6. Roshni, Babu S et al (2017) Design and fabrication of an E-shaped wearable textile antenna on PVB-coated hydrophobic polyester fabric. *Smart Mater Struct* 26(10):105011
7. Joshi JG, Pattnaik SS, Devi S (2012) Metamaterial embedded wearable rectangular microstrip patch antenna. *Int J Antennas Propag*
8. Sharma S et al (2018) Microstrip E-shaped patch antenna for ISM band at 5.3 GHz frequency application. In: 2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE). IEEE
9. Schilingovski P et al (2017) Wearable antennas design for wireless communication. In: 2017 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS). IEEE
10. Aprizal M et al (2018) Flexible artificial magnetic conductor reflector for wearable antenna application. In: 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama). IEEE
11. Joler M, Boljkovac M (2018) A sleeve-badge circularly polarized textile antenna. *IEEE Trans Antennas Propag* 66(3):1576–1579
12. Saluja R et al (2008) Analysis of bluetooth patch antenna with different feeding techniques using simulation and optimization. In: 2008 international conference on recent advances in microwave theory and applications, IEEE
13. Roh J-S, Chi Y-S, Kang TJ (2010) Wearable textile antennas. *Int J Fashion Des Technol Educ* 3(3):135–153
14. Balanis CA (ed ) (2011) *Modern antenna handbook*. John Wiley & Sons
15. Garg R et al (2001) *Microstrip antenna design handbook*. Artech house
16. Samaras T, Koulougou A, Sahalos JN (2004) A note on the impedance variation with feed position of a rectangular microstrip-patch antenna. *IEEE Antennas Propag Mag* 46(2):90–92

# Digital Storytelling and ICTs for Education to Foster Sustainable Development



Komal Ramdey and Hasnain Bokhari

**Abstract** Despite the efforts made by India with its programmes like Sarva Shiksha Abhiyan (Education for all), education is still not accessible for many. COVID-19 has further worsened the situation, as nearly 375 million children in India are likely to face disruptions in education because of the school closures. The situation for girls seems even murkier. It is estimated that nearly 13 million girls will be pushed into child marriage, as they will drop out of the school. This paper makes an attempt to contextualize e-learning and digital storytelling to foster development supported via ICTs. The paper discusses a case of not-for-profit fund “Going to School” in India that employs ICT tools and design thinking to create visual stories to teach entrepreneurship to children. The paper argues that innovative pedagogy, multichannel ICTs and partnerships with different government and corporate organizations hold a strong potential to make education accessible to the wider audience.

**Keywords** E-learning · Digital storytelling · Multichannel ICTs · Sustainable development · Going to school · COVID-19 · Tech4Good · Education · Innovation

## 1 Introduction

Technological progress and communication tools continue to redefine our relationship with a new global economy that is “powered by technology, fuelled by information and driven by knowledge” [1]. Rapidly evolving networked society constantly aspire all the sectors to incessantly acquire and apply new skills. With the change in demand side of the industry, education sector (supply side) also has to undergo a lot of changes. The International Labour Organization (ILO) defines the requirements for education and training in the new global economy simply as “Basic Education for All”, “Core Work Skills for All” and “Lifelong Learning for All” [2]. Therefore, the

---

K. Ramdey (✉)

Willy Brandt School of Public Policy, University of Erfurt, Erfurt, Germany

H. Bokhari

University of Erfurt, Erfurt, Germany

e-mail: [hasnain.bokhari@uni-erfurt.de](mailto:hasnain.bokhari@uni-erfurt.de)



effective integration of ICTs in the education sector can create an enabling environment in reaching these goals. The COVID-19 pandemic has left around 214 million children skipping more than three quarters of their in-person learnings. Similarly, 888 million children continue to face disruptions to their education due to full and partial school closures [3]. These precedents have unfolded many structural and societal inequalities. In a country like India, where 270 million people still live below the poverty line [4], nearly 375 million children are likely to face losses in education and work productivity [5, 6].

Indian school system relies on old pedagogies and employs the teacher-centric methods in practice for centuries. With the nationwide lockdown, it became very difficult for the schools to remain functional. Several private schools have switched to the online learning supported via digital communication platforms. However, for online teaching to work, Internet connection along with the laptop, tablet or smartphone are the essential prerequisites. Similarly, considering the teacher-centric pedagogy, online teaching alone does not entirely ensure the learning curve. Coronavirus thus acts as a two-edged sword for the Indian education sector with impediments related to the ICT access and employment of modern pedagogical skills suited for the young learners. Therefore, for the government as well as policymakers within the education industry, coronavirus is a paradox of sorts where the state has to ensure the ubiquity and affordability of communication technology which is no longer a luxury rather a necessity to survive, including and especially the ones at the bottom of pyramid.

Hereinafter, this paper wishes to investigate the potential that ICTs offer to make education accessible to all. Alongside, the paper also intends to explore the case of digital storytelling as a possible method to be employed in online teaching. In this context, the paper discusses the work of an India-based NGO “Going to School” that has relied on digital storytelling as a possible tool to teach children particularly during the pandemic crisis. In so doing, the paper will also highlight the role collaborative efforts/public private partnership to achieve the goal of “Basic Education for all” of SDG 9—Industry, Innovation and Infrastructure and SDG 17—Partnerships for Goals to achieve SDG 4—Quality Education and SDG 8—Decent work and Economic Growth.

## **2 Digital Storytelling, India and ICTs for Fostering Development**

In past few decades, digital storytelling has emerged as a formidable teaching and learning method. With the rise of user-generated content over the Internet, digital storytelling is stretching different creative boundaries for both teachers and students alike. The central concept of digital storytelling is to employ computers by creative storytellers to innovatively convey the narrative. Every story goes through the traditional process of subject selection and research, preparing a storyboard and developing an interesting story [1]. India also has a long tradition of storytelling where

each region has established its own style and tradition. Numerous nomadic groups often referred to as “Bhaand”, “Kathavachak” or “Kaawadiyas” who moved from village to village to perform and narrate stories [2, 3]. Similarly, Indian mythological and religious traditions have seen prominent scholars narrating the scriptures [4]. Within the education sector, digital storytelling offers a unique and stimulating experience of cultural nuances and visual and information literacy. It encourages students to learn the value of emotional rhetoric and empowers them to think differently and communicate confidently. When the digital storytelling movement started in 1980s in Berkeley, the electronic equipment involved bulkier desktop computers, separate video and sound recording devices and interfacing hardware along or scanners to show images from photograph albums. However, the dramatic growth in the ICT sector in past two decades has equally influenced the storytelling. From hefty computers, the equipment required for digital narratives has continued to shrink and more readily available as the global ICT spending explains which is forecast to grow to 5.82 trillion USD in 2023 from 4.92 trillion USD in 2019 [5].

There has been an exponential increase in the Internet users worldwide from 1.1 billion in 2005 to 4 billion in 2019. 75% of the total world population now has an active mobile broadband subscription whereas over 57% of households have Internet access at home [6, 7]. Heeks has strongly advocated the investment in the ICT sector as the infrastructural development influences the broader development agenda with ICTs as a fundamental part of infrastructure [8]. Particularly, for the low-income countries, investment in ICTs is not only linked with job creation and fiscal advantages but it would also help in keeping the digital divides at lower levels. As such technology has played a significant role in both economic growth and several facets of social welfare [9]. As a lower-middle income country, India has the world’s second-largest pool of Internet users, about 600 million. Nevertheless, almost half of its population does not have an Internet access [10]. As the pandemic unleashed its wrath around the globe, India has been the worst hit. As lockdowns across the country have been in place, Internet connectivity served as a cautious reminder for several sectors to operate behind the screens. Education sector has been in tatters due to shutdowns causing disruptions for the 285 million young learners in India. Some private schools decided to provide an alternative by opting for Internet-based teaching [11]. Initially, it appeared as if Internet was a saviour of education sector. To a larger extent, the argument holds true; however, as the first year of coronavirus revealed, there are two major issues attached to Internet-based teaching, first structural and second more on cognitive lines. Online teaching only favoured those learners whose parents could afford Internet connectivity and its related gadgetry. Even for learners with all the bells and whistles, having Internet access and device literacy is not directly proportional to classroom environment. For students, online learning requires a different discipline, netiquettes whereas for teachers, the structure of schooling, teaching and assessment methodologies are entirely different. Several teachers complained of burnout as their instructional quality went down while they juggled tirelessly between in-person, remote students and screen-based teaching [12].

For India, that takes pride in achieving the UN’s Sustainable Development Goals (SDGs) by 2030 has been forced by the pandemic to reconsider its action-oriented

programmes to engage its youth in early education. The need to invest in education has increased more especially after the pandemic, as the return to normalcy remains unclear. With the limited resources, especially with the government schools in a state like Bihar, Uttar Pradesh and Rajasthan, abiding to UN's transformative 2030 agenda for SDGs of "Leaving No One Behind" seems challenging. The accessibility of education has become very difficult for the low-income groups. While the coronavirus exposed several vulnerabilities in governance, state and social structures, it also provided a small window of opportunity to technology firms and civil society organizations to work in parallels and coordinate with the government in its efforts to provide services thereby also meeting UN SDGs such as Goal 9 (Industry, Innovation and Infrastructure) complimenting SDG 17 (Partnerships for Goals) to achieve SDG 4 (Quality Education) and SDG 8 (Decent work and Economic Growth). Particularly in the education sector, a number of civil society organizations have come forward to compliment Indian government in its effort to provide education [13, 14]. New Delhi-based non-profit education fund Going to School (GTS) which has its roots in India since 2003 has also come forward to support the education sector by forming various partnerships. GTS is currently exploring the amalgamation of traditional storytelling and learning and offer it as a part of curriculum. GTS is relying on digital storytelling and Internet as a medium to scale and making e-learning accessible to children during the pandemic. The following section explores the work of GTS and its effort in reducing the possible negative impact of COVID-19 on education and learning.

### **3 Case Study—Digital Storytelling and Partnerships to Make Education Accessible for All “Going to School”**

Despite India's significant progress in improving the literacy rates, India is still home to 301 million illiterate people, 59% of which are the women [15]. As the schools stand deserted and the education system facing its worst disruption in India due to pandemic, there is a strong fear that 13 million girls may be pushed into child marriage because they will be dropping out of school [16]. When “Going to School” (GTS) was founded in 2003, it aimed at working with schools across India to teach life skills to the poorest and most marginalized children. The organization first started its operations in Indian state of Bihar and later scaled to other states. The primary objective of GTS was to provide livelihood opportunities to young learners and promote transition to and retention in the secondary school [17]. One of GTS's core focus areas is young women's retention in schools and transition to equitable work or some enterprise. Till March 2019, GTS ran its operations across different states in India; however, the nationwide lockdown forced GTS to come up with out-of-box strategies and solutions. The situation called for a holistic approach with new modalities to make education available and accessible. It is then when GTS decided

to go back to the deep-rooted tradition of storytelling to the school children over multiple communication channels by forming partnerships for a better outreach.

### ***3.1 GTS Model and Approach***

GTS's pedagogical approach is based on interaction and visuals while the instructional and educational aspects are focused on socio-economic issues. These life challenges are presented via graphical novels especially tailored towards young learners offering design thinking that encourages problem-solving skills instead of traditional rote learning. GTS develop graphic novels for children that feature young women in new roles who are shown either establishing their own enterprises or taking up on challenges such as climate change. The illustrations are depicted as an action-packed adventure that motivate children for developing livelihood skills such as branding guidelines for setting up this enterprise, logo or a website design or procedures for developing sustainable packaging. Similarly, the graphic novels are also tailored to inspire and aid project-based learning for instance commonly known project steps that an enterprise goes through in a calendar year with the help of a detailed bullet journal. It is interesting to note how these visual stories bring this design thinking from paper to practice by encouraging young learners to carry out a four-to-five-day action project that takes place out of the classroom during which young people in a team of four interview for instance, beekeepers or organic farmers and enquire about how their businesses work or where their interviewees live. Based on their research, young learners submit their plans at the end of four-day exercises. GTS also prepare video versions of these stories for the educators. These programmes are run particularly in low-income government schools through partnerships.

### ***3.2 Technology, Innovation and Impact***

In the post-March 2020 scenario of nationwide lockdown in India, government schools were the worst hit as they, unlike few private schools, could not afford the online teaching. Bihar government and GTS collaborated with each other to make their storytelling programmes accessible to the wider audience and through multiple channels. In this context, Bihar government tasked the national television to air classroom curriculum on their free-to-air channels in Bihar and across the country. Some of the projects undertaken during the initial months of lockdown were focused on teaching livelihood skills such as entrepreneurship whereas the content was provided in multiple languages for a wider outreach [18].

Table 1 provides a brief overview of these projects along with the content, impact and outreach. GTS approach seemed very innovative as it employed both traditional and modern ICTs to increase the capabilities of both educators and learners. The work done by GTS during lockdown seems very impactful that is equally influenced

**Table 1** GTS projects: content, outreach and impact

Project	Content	Outreach	Impact
School TV	DIY series on entrepreneurial skills	DD Rajasthan DD Bihar YouTube	200 +episodes reaching 20 million children
Map of me (MOM)	26 steps to become an entrepreneur 30 days telecast 90 episodes of MoM TV	Mapofme.in Metros targeted (Delhi, Mumbai, Bangalore) Translation in four languages	16,000+young people reached
Young entrepreneur's skill show (YESS)	572 h of entrepreneurial skill and learning 5 weeks of telecast	DD Bihar	15,000+dial-ins for online quiz 12,00,000 students and teachers reached App downloads across 190 districts

by the various partnerships invited for collaboration. This model could be considered as a pilot project to further explore the potential of Tech4Good in education sector.

### 3.3 Analysis of Issues and Challenges

As a not-for-profit organization, GTS seems to have made a notable impact in making education accessible during pandemic. GTS could only reach out to these many students because of the partnerships it established with district governments and corporate funders thereby contributing to the SDG 17. However, there is no evidence on how many learners were able to start their own businesses by attending GTS's lessons. On the other hand, GTS continued to work during the pandemic and innovated its operations by establishing more partnerships with organizations such as Ikea Foundation, British Asian Trust, Oracle, Echidna Giving or Population Council. As outreach is the first step towards accessibility, it would be interesting to see the success rate with an enhanced focus on impactful stories. GTS primarily relies on corporate social responsibility (CSR) funding for its operations, and such a journey may well be difficult with several hindrances along the way restricting or hampering such an innovative work. COVID-19 has also posed actual implications on the CSR funding, and it is estimated that CSR activities are likely to drop by 30–60% [19]. Therefore, organizations like GTS may have to explore other and alternative sources of income to sustain their programmes. Similarly, the new Foreign Contribution (Regulation) Act (FCRA) 2020, which disallows subgranting, also restricts the functioning of many NGOs. The act has also been critically viewed by the United Nations Commission for Human Rights [20].

## 4 Conclusion

Storytelling has been recognized as an important pedagogical approach for enhancing student engagement [21]. Within the ICT perspective, digital storytelling provides strong foundation and cognitive learning effect to students, particularly ones from a marginal background [22]. International aid and development organizations such as UNICEF or UNESCO are promoting storytelling as an important approach towards inclusion in sustainable development context [23]. The case presented in this paper provides lessons on multiple fronts. The impact created by digital storytelling in comparison to the frontal teaching reflects that ICTs use within education sector can further contribute to achieve the sustainable development goals. It also shows that for an economically and linguistically diverse country like India, e-learning requires a revised approach towards Internet as a singular channel particularly for households who do not even have access to television. GTS's work through partnerships is a great example of involving the strengths of CSOs for community outreach and corporates for CSR funds and expertise in technology. Coronavirus also serves as an urgent reminder that how within the parameters of revised FCRA rules an enabling environment for CSOs can be created to strengthen their work in supporting the government. India has made a phenomenal progress in lowering the digital divide, nevertheless, inclusive policies which support the implementation of such progressive programmes need to be thought through. India, for instance, is all set to implement its new National Education Policy, 2020 which includes an integration of coding and increased exposure to technologies. Such ambitious goals would immensely benefit from a more robust approach towards education accessibility as well as increased spending in education budget [24]. Digital storytelling along with the Partnerships for Goals (SDG 17) presents a unique opportunity for organizations like GTS and other CSOs, and this model could be replicated in other states of India as well, not as a replacement to the teacher–student classroom method but as an addition to make education inclusive and accessible to even those who are not seen or heard.

## References

1. Robin BR (2008) Digital storytelling: a powerful technology tool for the 21st century classroom. *Theory Pract* 47(3):220–228. <https://doi.org/10.1080/00405840802153916>
2. The Hindu An ancient art of story-telling. [Online]. Available: <https://www.thehindu.com/features/friday-review/history-and-culture/an-ancient-art-of-storytelling/article5371732.ece>. Accessed 26 Mar 2021
3. All Good Tales Storytelling traditions across the world: India. [Online]. Available: <https://allgoodtales.com/storytelling-traditions-across-the-world-india/>. Accessed 05 April 2021
4. The Statesman Telling the tale. [Online]. Available: <https://www.thestatesman.com/world/telling-the-tale-2-20286.html>. Accessed 06 Mar 2021
5. Statista Research Department Global ICT spending 2016–2023. [Online]. Available: <https://www.statista.com/statistics/946785/worldwide-ict-spending/>. Accessed 05 April 2021
6. ITU Statistics. [Online]. Available: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>. Accessed 05 Mar 2021

7. ITU Digital trends in Europe 2021: ICT trends and developments in Europe, 2017–2020. [Online]. Available: [https://www.itu.int/en/ITU-D/Conferences/WTDC/WTDC21/Documents/RPM/EUR/Digital-Trends\\_Europe-E.pdf](https://www.itu.int/en/ITU-D/Conferences/WTDC/WTDC21/Documents/RPM/EUR/Digital-Trends_Europe-E.pdf)
8. Heeks R (2010) Do Information and Communication Technologies (ICTs) contribute to development? *J Int Dev* 22(5):625–640. <https://doi.org/10.1002/jid.1716>
9. World Bank Global economic prospects 2008: technology diffusion in the developing world. [Online]. Available: <https://elibrary.worldbank.org/doi/abs/https://doi.org/10.1596/978-0-8213-7365-1>. Accessed 24 Mar 2021
10. The Print As digital divide widens, India risks losing a generation to pandemic disruption. [Online]. Available: <https://theprint.in/india/education/as-digital-divide-widens-india-risks-losing-a-generation-to-pandemic-disruption/568394/>. Accessed 24 Mar 2021
11. Gulf News How COVID-19 has left India's education sector in tatters. [Online]. Available: <https://gulfnews.com/world/asia/india/how-covid-19-has-left-indias-education-sector-in-tatters-1.78709688>. Accessed 24 Mar 2021
12. The New York Times Teaching in the pandemic: this is not sustainable. [Online]. Available: <https://www.nytimes.com/2020/11/30/us/teachers-remote-learning-burnout.html>. Accessed 24 April 2021
13. India Today How this NGO is helping orphaned underprivileged kids continue their education in Covid-19 lockdown. [Online]. Available: <https://www.indiatoday.in/education-today/feature/story/how-this-ngo-is-helping-orphaned-underprivileged-kids-continue-their-education-in-covid-19-lockdown-1737659-2020-11-03>. Accessed 24 Feb 2021
14. Pratham Connecting to children and communities in a time of crisis. [Online]. Available: <https://www.pratham.org/2020/04/01/connecting-to-children-communities-in-a-time-of-crisis-pratham-initiative-for-fun-learning/>. Accessed 24 Mar 2021
15. Observer Research Foundation Literacy in India: the gender and age dimension. [Online]. Available: [https://www.orfonline.org/wp-content/uploads/2019/10/ORF\\_IssueBrief\\_322\\_Literacy-Gender-Age.pdf](https://www.orfonline.org/wp-content/uploads/2019/10/ORF_IssueBrief_322_Literacy-Gender-Age.pdf). Accessed 24 Mar 2021
16. World Vision Covid 19 aftershocks. [Online]. Available: [https://www.wvi.org/sites/default/files/2020-05/AftershocksFINALVERSION\\_0.pdf](https://www.wvi.org/sites/default/files/2020-05/AftershocksFINALVERSION_0.pdf). Accessed 24 Mar 2021
17. MacArthur Foundation Going to school fund. [Online]. Available: <https://www.macfound.org/grantee/going-to-school-fund-44992/>. Accessed 12 Mar 2021
18. Going to School Going to school in India: annual report 2020. [Online]. Available: <https://www.goingtoschool.com/pdfs/Going-to-School-Annual-Report-2020.pdf>. Accessed 12 Mar 2021
19. AVPN The impact of COVID-19 on CSR funding for Indian NGOs. [Online]. Available: <https://avpn.asia/blog/the-impact-of-covid-19-on-csr-funding-for-indian-ngos/>. Accessed 13 Mar 2021
20. The Scroll FCRA: centre makes rules stringent for NGOs, those participating in politics to receive funds. [Online]. Available: <https://scroll.in/latest/978321/centre-makes-rules-stricter-for-ngos-and-those-participating-in-politics-to-receive-foreign-funding>. Accessed 12 Mar 2021
21. Neal L (2001) Storytelling at a distance. *eLearn Mag* 2001(5):4
22. UNESCO Inclusive Policy Lab Digital story telling as an innovative pedagogical tool. [Online]. Available: <https://en.unesco.org/inclusivpolicylab/e-teams/digital-story-telling-innovative-pedagogical-tool>. Accessed 2021/15/03
23. UNICEF Sharing African stories in the digital age. [Online]. Available: <https://www.unicef.org/southafrica/stories/sharing-african-stories-digital-age>. Accessed 2021/15/03
24. The Economic Times education sector budget 2021: 6% cut in allocation for education sector. [Online]. Available: <https://economictimes.indiatimes.com/industry/services/education/budget-2021-6-cut-in-allocation-for-education-sector/articleshow/80640500.cms?from=mdr>. Accessed 2021/03/12

# A Study on IoT-Based Health Monitoring Systems for Nurse Assistance



Anju Pratap, Nithin Prince John, Greshma Johnson, Hanna Susan George, S. Devika, Anjal A. Jose, Kesia Sam, San Jos Varghese, Stevin Mathew Suban, and Salves Varghese

**Abstract** Today, healthcare monitoring devices are deployed in health management services to ensure patient safety and to reduce the burden of caregivers and health workers. This study mainly focuses on automatic healthcare systems using IoT that support nursing needs in the public health sector. Nurses work extended hours without proper rest and are more prone to diseases due to contact with patients. So, we can reduce the burden of nurses using these devices. The main objective of this study is to know more about alerting nurses when the IV drip underflows, detecting oxygen saturation level with a pulse oximeter, ECG system for continuous cardiovascular

---

A. Pratap (✉) · N. P. John · G. Johnson · H. S. George · S. Devika · A. A. Jose · K. Sam · S. J. Varghese · S. M. Suban · S. Varghese  
Saintgits College of Engineering, Kottayam, India  
e-mail: [anju.pratap@saintgits.org](mailto:anju.pratap@saintgits.org)

N. P. John  
e-mail: [nithin.pj@saintgits.org](mailto:nithin.pj@saintgits.org)

G. Johnson  
e-mail: [greshma.j1721@saintgits.org](mailto:greshma.j1721@saintgits.org)

H. S. George  
e-mail: [hanna.sg1721@saintgits.org](mailto:hanna.sg1721@saintgits.org)

S. Devika  
e-mail: [devika.s1721@saintgits.org](mailto:devika.s1721@saintgits.org)

A. A. Jose  
e-mail: [anjal.aj1721@saintgits.org](mailto:anjal.aj1721@saintgits.org)

K. Sam  
e-mail: [kesia.s1721@saintgits.org](mailto:kesia.s1721@saintgits.org)

S. J. Varghese  
e-mail: [san.jv1721@saintgits.org](mailto:san.jv1721@saintgits.org)

S. M. Suban  
e-mail: [stevin.ms1721@saintgits.org](mailto:stevin.ms1721@saintgits.org)

S. Varghese  
e-mail: [salves.v1721@saintgits.org](mailto:salves.v1721@saintgits.org)



health monitoring, etc. The contribution of this paper lies in the study of different health monitoring systems that helps the nurses to do their work with ease.

**Keywords** ECG · Healthcare · Internet of things · IV · Monitoring · Pulse oximeter · Sensor

## 1 Introduction

In today's world, health workers play a critical role in our lives. The COVID-19 pandemic and its negative consequences drew further attention to the healthcare industry. This motivates us to be more creative and forward-thinking in our healthcare systems through the application of new technology. Development in information and communication technology (ICT) and the Internet of things (IoT) have brought in vast progress in the health sector and biomedical engineering. Considering the case of patients in the intensive care unit (ICU), it is essential to look after the patient's condition either by an attendant or duty nurse allotted to that bed. Manual negligence in such a scenario may lead to the death of a patient. When it comes to the healthcare industry, nothing is more valuable than a patient's health. In such a scenario, the quality of treatment provided matters, not the number of patients or money turnover. A nurse assistance system is a system that allows nurses to do less stressful work. Implementing such monotonous and stressful work is easy. By combining sensors with IoT, such a device will serve as a viable alternative to the current situation. It is an interconnected network of sensors linked with the Internet to eliminate the day-to-day difficulties that the nurses face. It will award them with the extra time that they can use effectively for other things, which in turn helps them to work more efficiently. Patient data transfer is also another vital functionality that will be an efficient alternative to the current scenario. In Sect. 2, IoT in healthcare is discussed. Section 3 presents conclusion.

## 2 IoT in Healthcare

When connected to the Internet, IoT has unveiled a world of possibilities in medicine. This advancement allows current medical devices to gather invaluable additional data, activate isolated cases, and give patients more control over their lives and care [1]. It allows health professionals to be more attentive and interact more efficiently with patients. Sensor-equipped IoT devices help track the position of medical equipment such as wheelchairs, defibrillators, nebulizers, oxygen pumps, and other monitoring devices in real time [2].

## **2.1 Pulse Oximeter Monitoring System**

Bakhri et al. [3] suggested a method for calculating the oxygen saturation percentage in the blood. Pulse oximetry is a non-invasive blood oxygen saturation control technique. The percentage of total hemoglobin saturation is used to measure oxygen saturation in the blood (SpO<sub>2</sub>). The various wavelengths of red light (660 nm) and infrared light (940 nm) captured by the light sensor (IC MAX30100) at the tip of the index finger after moving through vessels and capillaries are used in this article. It was also shown that the error accuracy is equal to that of consumer products although the cost is lower. The primary goal of this paper is to show how to use the Raspberry Pi 3 platform to perform low-cost, high-resolution oximetry. The instruments produced in this study have the following advantages: simple to build, relatively less expensive, and simple to use for short-term data collection and calculation. The drawbacks are that it is not mobile and that it is relatively impossible to take constant measurements since the sensor is not attached to the tip of the finger but instead touches it.

Von Chong et al. [4] suggested a method in which a single LED was used to eliminate the errors caused by increased melanin or nail polish. Traditional pulse oximeters depend on the difference in absorption coefficients between hemoglobin and deoxyhemoglobin to estimate blood oxygen saturation non-invasively using two LEDs at various wavelengths and a photodiode. Low oxygen saturation or wavelength transmission, on the other hand, changes the contrast to capture the expected light on both LEDs. This paper suggests a new approach that combines a single LED and a Buried Quad Junction photodetector, i.e., a multispectral sensor, to solve this issue. They minimize the modifying effects of inaccuracy by using a single LED for this basic alteration of the pulse oximetry theory. In vivo analyses reveal that all three precision measures of the proposed device fell below the commercial tolerance of 4% SpO<sub>2</sub>.

## **2.2 Intravenous Infusion System**

Rao and Supriya [5] have created a new specification for IV bottles as well as a liquid level and drop count measurement device based on a wireless sensor network. Texas instruments CC3200, sensors like ultrasonic sensor HC-SR04, and LM35 temperature sensor, and GSM SIM900A are among the system's components. It also has the benefit of being able to watch the status from anywhere using Wi-Fi and cellular networks. As a result, the device eliminates potential risks to the patient such as blood loss and blood backflow due to nursing incompetence.

**Table 1** Tabulation of the reviewed health monitoring systems

Year	Authors	Features	System overview
2020	Bakhri et al. [3]	It is not a handheld device and it does not have any batteries	The low-cost and efficient oximetry system is built using the IC MAX30100 and the Raspberry Pi 3
2019	Von Chong et al. [4]	Inaccuracy due to a single LED is corrected	A single LED is used, as well as a buried quad junction photodetector (multispectral sensor)
2020	Rao and Supriya [5]	Possible danger to patients can be overcome. Cost of the sensors used are expensive	Uses Texas Instruments CC3200, Ultrasonic sensor HC-SR04, LM35 temperature sensor, and GSM SIM900A
2020	Xu [6]	Reduced cost than traditional ECG machines	Uses ECG sensors, Arduino, Android phones, Bluetooth and cloud servers with the proposed system

### 2.3 ECG Monitoring System

For continuous cardiovascular health surveillance, Xu [6] suggested an IoT-aided electrocardiogram (ECG) monitoring system with safe data transmission. The device includes sensors for monitoring ECG, Arduino, Android-based phones, Bluetooth networking, and cloud servers (Table 1).

## 3 Conclusion

The use of IoT in various health monitoring schemes is addressed in this study. There is still space for further research into how IoT is used in all aspects of medicine. Early detection of any health condition will assist the patient in taking the necessary emergency steps to save his or her life. IoT health management devices will track patients in real time and notify them if something is wrong. In the present situation, the nursing community is the main service provider. Nurses can benefit from an IoT-enabled device that can provide timely warnings by reducing their manual work. However, IoT networks must provide the tools to ensure that confidential data is properly secured. In addition, the sensors used must be compact enough to be conveniently installed in a number of applications. Finally, using a range of in-depth reading and learning equipment will improve the precision and robustness of programs. The idea of a smart health management system based on IoT structures is a revolutionary contribution to medical science that can help to minimize health problems and prevent unnecessary deaths. A health monitoring system can be implemented using informatics based on the methods described above. The development

of a technology-enabled portable assist device would aid them in making decisions and alerting doctors while they were away.

## References

1. Econsultancy. <https://econsultancy.com/internet-of-things-healthcare/>
2. Wipro. <https://www.wipro.com/business-process/what-can-iot-do-for-healthcare/>
3. Bakhri S, Rosiana E, Saputra RC (2020) Design of low cost pulse oximetry based on Raspberry Pi. *J Phys Conf Ser* 1501:012003. <https://doi.org/10.1088/1742-6596/1501/1/012003>
4. Von Chong A, Terosiet M, Histace A, Romain O (2019) Towards a novel single-LED pulse oximeter based on a multispectral sensor for IoT applications. *Microelectron J* 88:128–136. <https://doi.org/10.1016/j.mejo.2018.03.005>
5. Rao KR, Supriya KE (2020) Design and development of IoT based intravenous infusion system. In: Hitendra Sarma T, Sankar V, Shaik R (eds) *Emerging trends in electrical, communications, and information technologies. Lecture notes in electrical engineering*, vol 569. Springer, Singapore. [https://doi.org/10.01007/978-981-13-8942-9\\_4](https://doi.org/10.01007/978-981-13-8942-9_4)
6. Xu G (2020) IoT-assisted ECG monitoring framework with secure data transmission for health care applications. *IEEE Access* 8:74586–74594. <https://doi.org/10.1109/access.2020.2988059>

# A Survey on Diabetes Mellitus Prediction Using Machine Learning Algorithms



Rashmi Srivastava and Rajendra Kumar Dwivedi

**Abstract** Nowadays machine learning has several applications in the health sector. Disease prediction is one of the important healthcare applications in which various diseases have been detected at a very early stage. The prediction is done on the standard datasets available in the public domain. Imbalance and null values are a very crucial problem in the datasets for predicting diseases in the health sector because each reading sample has very significant properties of the cause of disease. Diabetes is a kind of disease that affects a huge no. of citizens of our society. In this context, it is very much important to adopt modern technology like machine learning and deep learning. This paper presents a state of art for diabetes prediction using machine learning techniques. We also discussed the major findings of the survey in brief.

**Keywords** Diabetes · Artificial intelligence (AI) · Machine learning (ML) · Deep learning · Disease prediction · Feature selection · Health care

## 1 Introduction

What is diabetes? To answer this question, we have to understand the role of glucose levels in our blood. Diabetes is the situation in which glucose exceeds/lacks a level in the blood. From the layman's understanding, glucose is important for producing energy when it breakdowns and if this breakdown is unable to be done then the level gets misbalanced. The cause is the inefficient production of insulin from the pancreas. This insulin is a very important type of enzyme that converts sugar into energy which we used to make our body function. Diabetes is an incurable disease but luckily it is a treatable disease. So that is why in today's time we have to take preventions and early predictions as soon as possible. Because diabetes has attracted the research area a lot, some missing features need to be found then only the correlation will be better to other segments of the body. We have taken PIMA (PIMA Indian diabetes dataset) from the University of California, UCI (Irvine) Repository. We have to compensate for the missing values in the given dataset by allowing some algorithm to predict, then

---

R. Srivastava (✉) · R. K. Dwivedi  
Department of IT, CA MMMUT, Gorakhpur, UP, India

also used some algorithm to balance the dataset class and reduce the class imbalance factor.

In this paper, Sect. 2 gives an overview of some data generated by international organizations and the impacts of diabetes on the world's population. Further, Sect. 3 gives us the information regarding the required criteria in diagnosing diabetes. Next Sect. 4 defines how machine learning involves in this study. Section 5 presents a survey on various existing researches in the related field of work and their limitations. Section 6 describes the findings of the survey. Finally, Sect. 7 provides the conclusion and future work.

## 2 Impact of Diabetes Disease in Today's Scenario

Diabetes is a metabolic disease having chronic properties which persist for a long time that occurs when the pancreas does not make a certain amount of insulin that the body needed. According to WHO, between 2000 and 2016, a 5% increment is being there in premature\_mortality\_rate from diabetes. In developed countries, the precocious mortality rate because of diabetes decreased or minimized from 2000 to 2010 but increased in 2010–2016 but in developing countries, rates increased across both the time. Recently, in 2014, 8.5% of adults or old people aged 18 years and older have diabetes. But, till 2016, diabetes was the undeviating cause of 1.6 million demise or deaths. The diabetic patients grew from 108 million in 1980 to 422 million in 2014. The global common occurrence of diabetes among elder or adults over 18 yrs of age grew from 4.70% in 1980 to 8.50% in 2014 (WHO official site).

## 3 Criteria for Diagnosing Diabetes

According to the University of Michigan, some criteria should be met like, have symptoms of diabetes and a blood sugar level. This sugar test is been done after not taking or drinking anything but water for 8 h, having a 2 h oral-glucose tolerance test (OGTT) result  $\geq 200$  mg/dL, have a hemoglobin A1c that is 6.50% or above. These are the only criteria that meet the requirement. But two tests are needed to ensure the possibility of diabetes. According to WHO's recommendation, the HbA1C test has evolved as the most popular method of diagnosing for insulin shortage. It measures part of the RBCs because glucose in the blood attaches "t" part of the RBCs. This test gives a good report of having average blood glucose 2–3 months before. But still, they have some limitations like they cannot be used with people having age less than 22 or if anyone has been suspected of type 1 diabetes already or in pregnancy. These are the most important prohibition related to the HbA1C Blood test. The second test related to diagnosing diabetes via FBG test, that is taken using also via blood in the period of fasting of at least 8 h. A sample is taken early the next day of the fasting day. To confirm a diagnosis, repetition of testing after a certain interval is required.

This test has many phases of diagnosis so that we should take really good care so that we easily pass all the phases.

## 4 Machine Learning Techniques Applicable for Disease Prediction

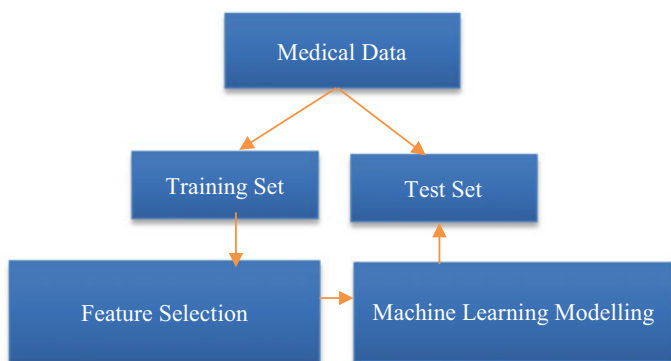
Machine learning is a form of artificial intelligence that enables multiple datasets for robust interrogation to signify past undiscovered sequences and relationships in the data. Machine learning is used these days in various types of applications, viz. outlier detection, anomaly detection, finding suitable advertisements, and disease prediction [1–4]. It is the main demand by the medical units. Medical units have a large amount of data that is used as a weapon against unknown disease predictions and cures also. These large datasets will be important to achieve full integration of machine learning for gaining an enhanced version of the treatment. Medical colleges and hospitals store the data about large amount patients and also sent to different medical research organization like aims, who, etc., for estimating the situation of citizens' health and disease procurement.

Here, machine learning models are very helpful because big data means the data about data are very complex to handle deep learning, and machine learning is always ready to reduce this complexity. Medical science has a variety of disease problem statements that need to be predicted likewise especially those diseases which are incurable but treatable like diabetes mellitus and its types (TYPE1, TYPE2). As we know, we have various types of machine learning algorithms. In the second situation where some pattern-finding situation involves then, we should use the clustering method in which utilizes the gathering of data points. Next situation where we have to predict true or false situations regarding questions asked, then we can go for a neural network and others will help. In those situations where a decision framework is required to build, then the decision tree will be fitted in demand and so on. Figure 1 presents medical usage in machine learning modeling, and Fig. 2 shows a situation-based classification of a machine learning algorithm in the medical field.

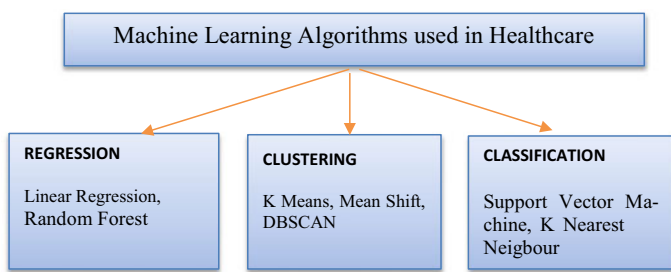
## 5 Literature Review

To discover a solution to the diabetes prediction problem, we have to review all previously done work to give more effective solutions. So this part of the paper is reviewing all related work as mentioned in Table 1.

Zhang et al. [5] proposed a technique that uses classification and regression tree (CART's) applied to tackle the problem. C—CLASSIFICATION A—AND R—REGRESSION T—TREE. Cart initialization means producing a decision tree based on training datasets. Cart pruning and optimization regression tree is sheared



**Fig. 1** Use of machine learning modeling with medical data



**Fig. 2** Situation-based classification of a machine learning algorithm in the medical field

according to some constraints (like max depth of the tree, min sample no. of the leaf, and node's min impurity). Kumar and Umatejaswi [6] suggested a system in which techniques used-KNN, data preprocessing, apply k-means, apply to classify algorithm to classify the patient's risk level of diabetes.

Zeng et al. [7] have given an idea that uses Synthesis Minority Oversampling technique (SMOTE) as an effective data sampling algorithm has been applied to analyze unbalanced medical data. This paper compared SMOOTH with eight classifiers SVM, BN, AdaBoost, K, C4.5, RBF NETWORK, LR, and LMT. Liu et al. [8] have given this idea to use fuzzy logic in context for predictions. This has fuzzy-based information decomposition (FID) that solves two problems—FID balances the train data by making additional samples for the non\_major class. These techniques have two steps—(1) weighing and (2) recovery.

Jamgade and Zade [9] proposed an idea to use multiple dimensions of data—in this paper, the authors combine the study of the structure and unstructured data in the medical sector that increase the assessment of prediction of diseases. So it handles the determination of chronic diseases in a particular region in a specific community. Ashiquzzaman et al. [10] have given a system of prediction which uses data overfitting in the predictive model as a dropdown technique. In this work, proposed work, a novel



**Table 1** A comparative study of the existing standard machine learning-based diabetes mellitus prediction schemes

Authors	Techniques used	Results	Limitations
Zhang et al. [5]	Support vector machine	78% accuracy	Complex task handling not possible
Kumar and Umatejaswi [6]	K-nearest neighbor	78.16% accuracy	Cannot give you unknown information from trained data
Zeng et al. [7]	Smote Tomek and comparison between 8 classifier	Results are better with Smote-Tomek rather than only smoke	Each has different values of result variants
Liu et al. [8]	Fuzzy logic	64.44% accuracy	Not capable like machine learning or neural network
Jamgade and Zade [9]	K-means	95% accuracy	No supervising element
Ashiquzzaman et al. [10]	Novel form of deep neural networks	88.41% accuracy	More time spending in interpreting and labeling class
Yahyaoui et al. [11]	SVM, DL, RF	83.67% accuracy	Clustering issue
Hasan et al. [12]	Ensemble learning (AB + XB) model	95.0% accuracy	Difficulty in interpretation
Aliberti et al. [13]	NL aggressive (NAR) and LSTM, neural network	Compared with three schemes (FNN, RNN, AR)	Fine results only for short-term prediction
Lee et al. [14]	LR and Naïve Bayes classifier	Females—73%–74% Males—68.6%–8.7%	FPG-level prediction main cause—diabetes
Sierra-Sosa et al. [15]	LDA, SVM, RNN	Above 90%	—
Li et al. [16]	Deep learning model	RMSE best among other algorithm	—
Sneha and Gangil [17]	Decision tree, random forest, SVM, NB	DT (98.00%), RF (98.02%), SVM (77.73%), NB (82.30%)	Able to map the attribute successfully from low to high dimensions only
Vizhi and Dash [18]	KNN, logistics regression, gradient boosting	77%	Clustering issues
Lee and Kim [19]	LR, HW (hyper-triglyceridemic waist)	IN MALE-HW strongest, type 2 diabetes IN FEMALE-HW is connected with type 2 diabetes	HW Phenotype, etc., individual measurement is little complicated
Georga et al. [20]	Tenfold cross-validation, SVR model	Data growth in the precision of both short-term and long-term prophecy	Availability of multivariable focused only

form of deep neural network has been used. Yahyaoui et al. [11] have given a system in which does a comparison-based study of machine learning and deep learning-based algorithms for prediction of diabetes is given. Random forest, SVM, and deep learning are used, but in this paper, random forest performs best among all three also used. Hasan et al. [12] suggested an ensemble technique for the prediction of diabetes in which AdaBoost and XgBoost combination has recovered the highest accuracy rate after comparing with various combinations of machine learning algorithms. Here, he presented a substructure for diabetes prediction where outlier, missing values, data standardization, feature selection, cross-validation, and different classifiers and multilayer perceptron (MLP) were used. Aliberti et al. [13] proposed an approach in which we check the prediction models use data of glucose signals of a huge and heterogeneous cohort of sufferer and then applied to infer future glucose levels values on a whole new patient. Two different types of solutions like NAR and LSTM networks are compared with three literature approaches (FNN, RNN, AR).

Lee et al. [14] studies aim to predict fasting plasma glucose taken and categorized by logistic regression and Naïve biased classifier. Sierra-Sosa et al. [15] maintain a very good idea of evaluation of diabetes on variation parameters, in diabetes an acute myocardial infarction was obtained by LDA, SVM gives the accuracy of 90+% and having inclusion of hospital admission for avoidable causes—LDA and SVM both with 92%, RNN gives 94.6%, so thus on variations, he studies the best classical machine learning techniques studies. Li et al. [16] presented a deep learning model that is mainly capable of prognosticating glucose levels with leading accuracy which is the main cause of having diabetes. Their model provides competency in form of effective prediction horizon (PHeff) with a minimal\_time\_lag in a simulated patient dataset and a real patient dataset also.

Sneha and Gangil [17] have designed a prediction algorithm which is machine learning and search for the optimal classifier to give the adjacent result. Predictive analysis is the technique for selecting attributes and then do the early detection of diabetes mellitus. The resulting value shows that the decision tree algorithm and the random forest have the highest specificity of 98.20% and 98.00%, for prediction. Vizhi and Dash [18] give a solution regarding suffering from diabetes by giving ideas via paper up with smart trigger and different machine learning algorithms. There is an uplifting scheme also given available for use of train-test split and K-fold cross-validation using the scikit-learn method.

Lee and Kim [19] used binary logistic regression (LR) to inspect statistically notable dissimilarity between normal subjects and type 2 diabetes using HW and individual anthropometric measurements. For more valid and genuine prediction results, two machine learning algorithms are used here, Naive Bayes (NB) and logistic regression, for evaluating the predictive power of various phenotypes. Georga [20] proposed the idea of having three variables that is glucose profile, plasma insulin concentration, and the arrival of meal-derived glucose in the systematic movement and energy expenditure. Six cases have taken of mentioned variables for investigation, and the proposed work is validating through the tenfold cross-validation and testing by the SVR model.

## 6 Major Findings of the Survey

By this survey, it can be said that there are different types of machine learning algorithms such as SVM, KNN, AdaBoost, decision tree, K-means, deep network, and random forest that can be used for diabetes prediction. A comparative study of these algorithms is presented in this survey. According to this survey, we can say that the K-means and  $(AX + XB)$  model [12] are used in ensemble learning. SVM, RNN, LDA, and RF show 90% or above accuracy level.

## 7 Conclusion and Future Work

Surveys are always important when data collection and its analysis becomes essential in targeting the group of people of some specific horizon. Here, we compared the significance of several machine learning algorithms used for diabetes prediction such as support vector machine (SVM), K-nearest-neighbor (KNN), AdaBoost, decision tree, K-means, deep network, random forest, recurrent neural network (RNN), and linear discriminant analysis (LDA). SVM gives the best accuracy among these algorithms. This survey adds some significant trends analysis under medical field came in front during over a while. In future, we can implement the optimized model for disease prediction by selecting the best machine learning algorithms.

## References

1. Dwivedi RK, Kumar R, Buyya R (2021) Gaussian distribution-based machine learning scheme for anomaly detection in healthcare sensor cloud. *Int J Cloud Appl Comput (IJCAC)* 11(1):52–72
2. Dwivedi RK, Rai AK, Kumar R (2020) A study on machine learning based anomaly detection approaches in wireless sensor network. In: 10th IEEE international conference on cloud computing, data science and engineering (confluence-2020), Amity University Noida, pp 200–205. <https://doi.org/10.1109/Confluence47617.2020.9058311>
3. Dwivedi RK, Rai AK, Kumar R (2020) Outlier detection in wireless sensor networks using machine learning techniques: a survey. In: IEEE international conference on electrical and electronics engineering (ICE3–2020), organized by MMMUT Gorakhpur and North Dakota State University, Fargo, USA, at MMMUT Gorakhpur, pp 316–321. <https://doi.org/10.1109/ICE348803.2020.9122865>
4. Dwivedi RK, Pandey S, Kumar R (2018) A study on machine learning approaches for outlier detection in wireless sensor network. In: 8th IEEE international conference on cloud computing, data science and engineering—confluence (ISBN 978-1-5386-1718-2), held at Amity University, Noida, India, pp 189–192. <https://doi.org/10.1109/CONFLUENCE.2018.8442992>
5. Zhang B, Wei Z, Ren J et al (2018) An empirical study on predicting blood pressure using classification and regression trees. *IEEE Access* 6:21758–21768
6. Kumar PS, Umatejaswi V (2017) Diagnosing diabetes using data mining techniques. *Int J Sci Res Publ* 7:705–709

7. Zeng M, Zou B, Wei F, Liu X, Wang L (2016) Effective prediction of three common diseases by combining SMOTE with Tomek links technique for imbalanced medical data. In: IEEE International Conference of Online Analysis and Computing Science (ICOACS), Chongqing, pp 225–228. <https://doi.org/10.1109/ICOACS.2016.7563084>
8. Liu S, Zhang J, Xiang Y et al (2017) Fuzzy-based information decomposition for incomplete and imbalanced data learning. *IEEE Trans Fuzzy Syst* 25(6):1476–1490
9. Jamgade AC, Zade SD (2019) Disease prediction using machine learning. e-ISSN: 2395-0056
10. Ashiquzzaman A et al (2018) Reduction of overfitting in diabetes prediction using deep learning neural network. In: Kim K, Kim H, Baek N (eds) *IT convergence and security 2017. Lecture notes in electrical engineering*, vol 449. Springer, Singapore. [https://doi.org/10.1007/978-981-10-6451-7\\_5](https://doi.org/10.1007/978-981-10-6451-7_5)
11. Yahyaoui A, Jamil A, Rasheed J, Yesiltepe M (2019) A decision support system for diabetes prediction using machine learning and deep learning techniques. In: 1st International informatics and software engineering conference (UBMYK), Ankara, Turkey, pp 1–4. <https://doi.org/10.1109/UBMYK48245.2019.8965556>
12. Hasan MK, Alam MA, Das D, Hossain E, Hasan M (2020) Diabetes prediction using ensemble of different machine learning classifiers. *IEEE Access* 8:76516–76531. <https://doi.org/10.1109/ACCESS.2020.2989857>
13. Aliberti A et al (2019) A multi-patient data-driven approach to blood glucose prediction. *IEEE Access* 7:69311–69325. <https://doi.org/10.1109/ACCESS.2019.2919184>
14. Lee BJ, Ku B, Nam J, Pham DD, Kim JY (2014) Prediction of fasting plasma glucose status using anthropometric measures for diagnosing type 2 diabetes. *IEEE J Biomed Health Inform* 18(2):555–561. <https://doi.org/10.1109/JBHI.2013.2264509>
15. Sierra-Sosa D et al (2019) Scalable healthcare assessment for diabetic patients using deep learning on multiple GPUs. *IEEE Trans Industr Inf* 15(10):5682–5689. <https://doi.org/10.1109/TII.2019.2919168>
16. Li K, Daniels J, Liu C, Herrero P, Georgiou P (2020) Convolutional recurrent neural networks for glucose prediction. *IEEE J Biomed Health Inform* 24(2):603–613. <https://doi.org/10.1109/JBHI.2019.2908488>
17. Sneha N, Gangil T (2019) Analysis of diabetes mellitus for early prediction using optimal features selection. *J Big Data* 6:13. <https://doi.org/10.1186/s40537-019-0175-6>
18. Vizhi K, Dash A (2020) Diabetes prediction using machine learning. *Int J Adv Sci Technol* 29(06):2842–2852. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/13795>
19. Lee BJ, Kim JY (2016) Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on machine learning. *IEEE J Biomed Health Inform* 20(1):39–46. <https://doi.org/10.1109/jbhi.2015.2396520>
20. Georga EI et al (2013) Multivariate prediction of subcutaneous glucose concentration in type 1 diabetes patients based on support vector regression. *IEEE J Biomed Health Inform* 17(1):71–81. <https://doi.org/10.1109/TITB.2012.2219876>

# DisasterNet: a Multi-label Disaster Aftermath Image Classification Model



Subham Singh, Sandeep Ghosh, Amit Maity, Bikash Chandra Bag,  
Chaitali Koley, and Hirak Kumar Maity

**Abstract** Natural disasters can occur without any warnings, after a disaster passes by, all that we are left with is destruction of flora and fauna along with a huge imbalance in the ecosystem of that particular region. A country's level of development can be judged very easily by its rate of recovery from such natural disasters. Hence, here we have created a classification model based on deep neural networks that can successfully mark the damaged areas using ground images of those areas, scraped from different social media platforms uploaded right after the disaster has occurred. As a result, this can accelerate the work of recovery very easily by marking the areas where the disaster has more impact than the others. We have used Google's MobileNetV2 and created a multi-output classification model optimized using Nadam optimizer and trained on a custom data set containing 2000 (approx.) different disaster aftermath multi-labelled images belonging to different classes. After successful completion of training, we have tested our model's efficiency on two custom data sets. Accuracy, precision, recall and F1-score were the metrics utilized for the evaluation of the test results.

**Keywords** Disaster image classification · Multi-label classification · Deep learning · Computer vision · MobileNetV2

## 1 Introduction

Disasters can be man made as well as natural. Man-made disasters can be controlled to some extent, but natural disasters are something that we cannot put brakes on.

---

S. Singh (✉) · S. Ghosh · A. Maity · H. K. Maity  
College of Engineering and Management, Kolaghat, West Bengal, India  
e-mail: [ece1722@cemk.ac.in](mailto:ece1722@cemk.ac.in)

A. Maity  
e-mail: [hmaity@cemk.ac.in](mailto:hmaity@cemk.ac.in)

B. C. Bag · C. Koley  
National Institute of Technology Mizoram, Chaltlang, Aizawl, India

Once the disaster is over, all that we are left with is destruction and various instances which depict the impact of the occurred disaster that took place.

Restoring everything back to the way it used to be is indeed a tough job that requires proper and structured planning. The first step of restoration is the identification of the affected areas. Identification can be done in various ways, from sending ground teams to predicting the impact of the disaster by taking aerial images. But, in both the mentioned cases there are few drawbacks, sending ground teams to various places can be both expensive and time taking, and at the same time, predicting the original impact with the help of aerial images can be inappropriate too.

With the boom in the digitization, if we can scrap images of a particular area from multiple social media platforms just after the disaster has occurred, we can make a rough estimate of the areas which got highly affected just by analysing those scraped images and act accordingly. For analysing those images, we have made a multi-output convolutional neural networks classification model that is trained on a custom data set containing different disaster aftermath multi-labelled images divided into multiple classes, with the help of which we can identify the affected areas in an efficient way.

The organization of the paper is as follows: Sect. 2 briefly describes the related work. The proposed architecture and parameter optimization are mentioned in Sect. 3. Section 4 describes about the data set and its processing. Results and discussions are mentioned in Sect. 5. Section 6 describes the scope of the future works.

## 2 Related Works

This section provides a brief literature review about existing methods and recent works on classification/evaluation of natural disaster aftermaths. In the recent years, deep learning has achieved unprecedented success in computer vision and computer-aided disaster management.

1. Mao et al. [1] used a subset of the LADI data set, also known as the low altitude disaster imagery data set and trained multiple convolutional neural networks (CNNs) to classify between flood and non-flood images. Their ResNet101 model achieved an accuracy of 79% and an F1-score of 79.6 which they considered as their best.
2. Chaudhuri et al. [2] proposed a method to identify if human body parts are present in earthquake-hit debris images. They collected images of earthquake-hit zones from the National Hazards Image Database of the National Centers for Environmental Information. They also crowdsourced similar types of images from Google Images and then trained a seven-layer CNN with ( 500) images which resulted in an overall accuracy of 83.18% and a F1-score of 83.89.
3. Rosebrock [3, 4] in an article described a CNN that can automatically detect natural disasters (earthquakes, floods, wildfires, cyclones/hurricanes) with up to 95% accuracy. He used a data set that was curated by a PyImageSearch reader,

Gautam Kumar, where he used Google Images to gather a total of 4,428 images belonging to four separate classes, i.e. cyclone/hurricane (928 images), earthquake (1350 images), flood (1073 images) and wildfire (1077 images).

3 Proposed Method

The method we propose is a supervised learning algorithm based on a convolutional neural network, where we try to optimally fit a universal function approximator, i.e.  $f : x \rightarrow y$ . We map disaster aftermath images ( $x$ ) to their respective multiple labels ( $y$ ). We then optimize the parameterized neural network comprising of multiple layers of different types having different weights and biases using an iterative optimization technique, i.e. gradient descent to minimize a loss function  $L = g(x, y, \odot)$  by tuning the ' $\odot$ ' parameter.

3.1 Network Architecture

The base network is MobileNetV2 [5] (Fig. 1a(1)) which was developed by Google and was pre-trained on the ImageNet data set, a large data set consisting of 1.4M images and 1000 classes. We use this network as a fixed feature extractor to extract meaningful features from our custom data set. We freeze the entire network and replace its 1000 class classification block with a global average pooling block (GAP)(Fig. 1a(2)) and our own multi-output classification block. The output from the last layer of the frozen network is a feature map ( $W, H, D$ ), where the first two dimensions represent the spatial dimensions and the last dimension is the number of feature maps. This feature map is passed through a GAP layer where for each feature map, the average value of the spatial region is taken along with the number of feature maps.

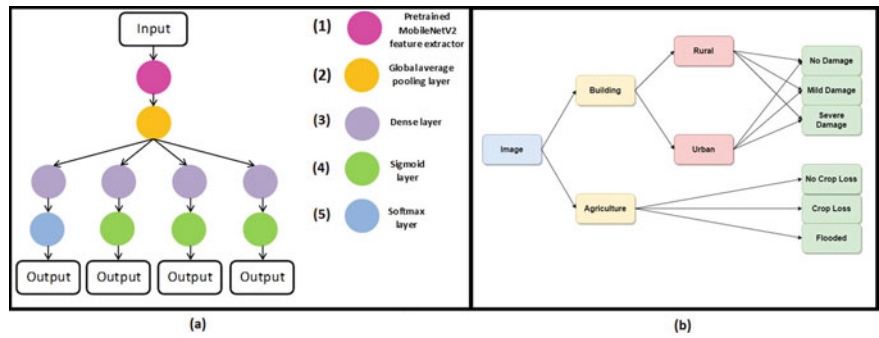


Fig. 1 Proposed network architecture

The output dimension hence is  $(1, 1, D)$ . As there is no parameter to optimize in the GAP layer, thus overfitting is avoided and the feature maps can easily be interpreted as categories confidence maps. Furthermore, GAP sums out the spatial information, thus it is a more robust alternative to spatial translations of the input. The output of this layer is then passed through *four* dense layers (Fig. 1a(3)) in parallel having 1024 neurons each. Once again all the outputs of the dense layers are connected with the final classification dense layer consisting of *two, three and four* neurons, respectively. The output layer with *two* neurons (Fig. 1a(4)) serves to classify the type of the image, i.e. whether the image contains buildings or an agricultural field. As this layer acts as a binary classifier, we have used sigmoid activation. Immediately next to the previously mentioned output layer, another dense layer with three neurons (Fig. 1a(5)) with softmax activation aims to classify whether the image belongs to an urban, rural or neither of these locations. Next, we have another dense layer with *four* neurons (Fig. 1a(5)) and a softmax activation which classifies whether the agricultural field has suffered crop loss, is flooded, no crop loss or neither of these. Following this layer, we have another identical dense output layer (Fig. 1a(4)) which classifies whether the building is mildly or severely damaged or there is no damage at all. The proposed network architecture is shown in Fig. 1a. It receives a three channel (RGB) input image of dimension  $(256, 256, 3)$ .

### 3.2 Parameter Optimization

The optimizer that we have used here is Nadam Wang et al. [10]. It is a combination of Nesterov accelerated gradient (NAG) [6] along with adaptive moment estimation (Adam), [7], Zhang [8], Bock et al. [9]. As we know, Adam is computationally efficient and has very little memory requirement and NAG works better than standard momentum as it can optimize the descent faster because it calculates the gradient with respect to the future step rather than the current step [11–13].

Thus, as a result, in Nadam, the learning process is accelerated by summing up the exponential decay of the moving averages for the previous as well as the current gradient.

## 4 Data Set

The network is trained on a 16-GB Nvidia P100 GPU paired up with an Intel Xeon CPU and 12GB RAM. The deep learning framework used is Keras with TensorFlow 2.X as the back end.



## 4.1 Data Collection

We proposed a custom data set containing disaster aftermath images. The dataset contains ( 1900) images crowdsourced from various platforms like Google Images, Flickr, Bing, etc. Our data set contains four interconnected classes, such that each downloaded image is labelled manually on the following basis.

- Type (building or agriculture)—This is our super class which decides that the image contains buildings or agricultural fields.
- Location (urban, rural, none)—This is our second interconnected class. Here the images are labelled on the basis of its location. If the image is labelled as building in the super class, then it gets further labelled as rural or urban depending on its location, and if it is classified as agriculture in the super class, then it gets labelled as none in this subclass.
- Building damage severity (no damage, mild damage, severe damage, none)—This is another subclass, here if the image is labelled as building → urban or rural, then it gets further labelled as no damage, mild damage or severe damage on the basis of visual judgement over the crowd sourced data. And, if it belongs to agriculture → none, then the image again gets labelled as none.
- Agriculture field damage severity (no crop loss, crop loss, flooded)—This is the last subclass of the data set, if the image is labelled as agriculture → none → none, then it further gets labelled as crop loss, no crop loss or flooded based on the image. And, if the image is labelled as building → urban or rural → no damage, mild damage, severe damage, then it gets labelled as none.

On a more elaborative way, each image was labelled considering the structure shown in Fig. 1b.

## 4.2 Scaling and Grouping of Data

We have scrapped the images from different data platforms like Google Images, Bing and then preprocessed those images and labelled them into different classes and subclasses based on our visual thresholding. For examples, some of the images along with the classes they are associated with are shown in Fig. 2.

## 4.3 Data Preprocessing

Each image in the data set is resized to (256, 256, 3). Rescaling is applied on the fly to normalize the image pixel values to be between [0, 1]. As the data set is relatively small, random augmentations like random flips, zooming and rotations with random changes in lightning and channel shifts are applied to manage the images [building,

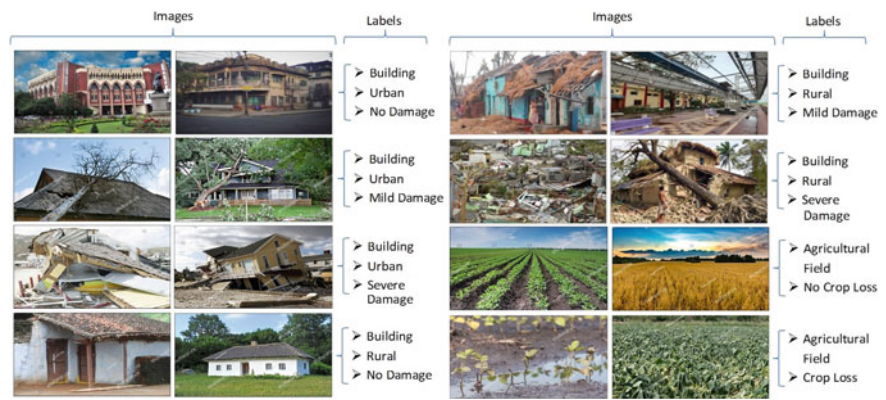


Fig. 2 Images along with classes

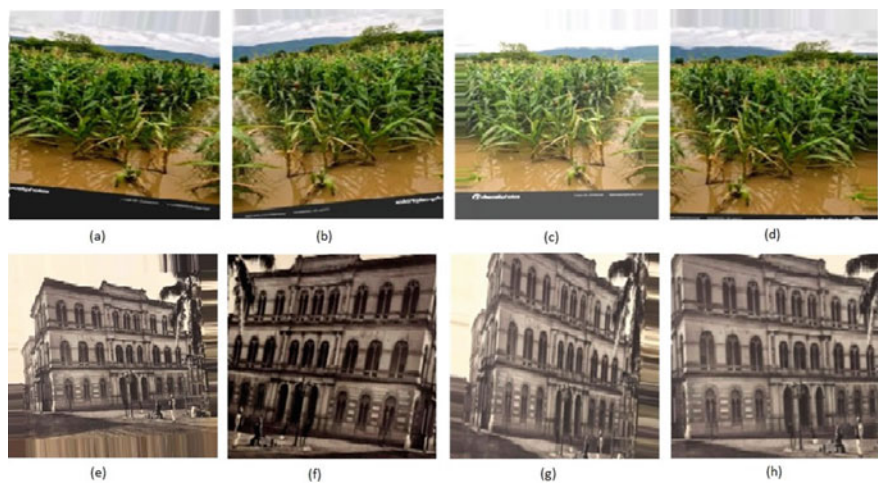
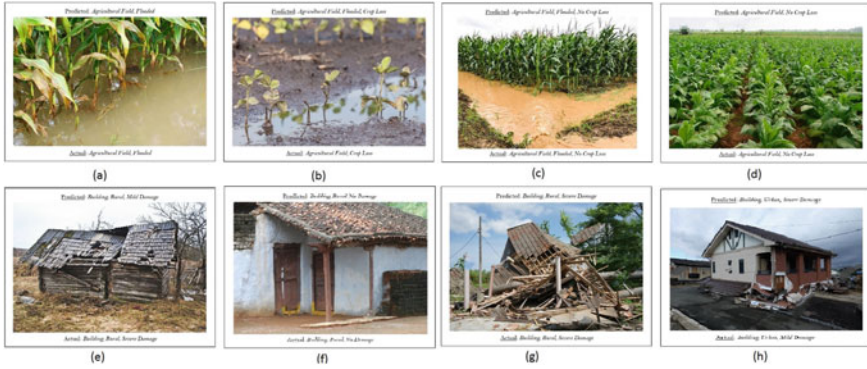


Fig. 3 Sample of some augmented Image

rural, no damage] [building, rural, severe damage] into equal image sizes. Some of the augmented images are shown in Fig. 3.

5 Results and Discussion

This section consists of a detailed report on the final outcome of our extensive research and analysis along with a discussion on the further scopes of development in the near future.



**Fig. 4** Predicted output images using proposed model

**Training and Validation:** We have compiled our model using Nadam optimizer and used binary cross-entropy along with sigmoid activation function for binary classification, and categorical cross-entropy along with softmax activation was used for classifying distributions consisting more than two classes.

$$Precision = (TP)/((TP + FP)) \quad (1)$$

$$Recall = (TP)/((TP + FN)) \quad (2)$$

$$F1 - Score = ((2PrecRec))/((Prec + Rec)) \quad (3)$$

where TP, FP and FN are the numbers of true positives, false positives and false negatives. The accuracy score is not used as an evaluation metric because the true negatives outweigh the true positives significantly and are therefore easy to predict. Therefore, the number of true negatives does not correctly reflect the quality of the prediction. Also some predicted images from our proposed model are shown in Fig. 4.

**Comparative Studies:** As our data set is custom and not related to any of the available open source data sets, we cannot directly compare our proposed method with any other methods, but we are providing a remote comparison based on relevant/near-relevant previous works in the domain of ground level disaster aftermath image classification in Table 1.

## 6 Conclusion and Scope of Future Work

The proposed method is based on a lightweight MobileNetV2 architecture with a custom classification block which was trained with a custom, crowdsourced multi-label data set which achieved descent performance levels setting a new benchmark for the

**Table 1** Comparisons of the performance measured

Works	Data set	No. of images	Architecture	Class	Accuracy	Precision	Recall	F1-score	Model size
[1]	Flood versus non-flood images derived from LADI data set	2000	ResNet101	NA	79	79.5	79.9	79.6	> 163
[2]	Human body parts from debris images extracted from the NHI database and Google Images	500	Seven-layer CNN	NA	83.1	85.87	82.6	83.89	NA
[3]	Earthquake flood wildfire cyclone	4428	VGG16	Cyclone earthquake flood wildfire	95.0	99 96 90 96	97 93 94 97	98 95 92 96	NA
Proposed method				Agro field	93.0	82.5	99.7	90	62
				Building	93.0	99.2	88.5	93.5	
				Urban	91.0	99	85	91.4	
				Village	91.0	82	90	85.8	
				N/A	91.0	91	96	93.4	
	Custom Crowdsourced Multi-label Data set	1900	Mobile NetV2	Crop loss	88.5	99	99	99	
				Flooded	88.5	57	57	57	
				No crop loss	88.5	67	90	76.8	
				N/A	88.5	98	97	97.5	
				No damage	80.0	98.9	70	81.9	
				Mild damage	80.0	68.7	75	71.7	
				Severe damage	80.0	99	62	76.2	
				N/A	80.0	72	99	83.3	

task of disaster response and assessment as disaster-hit zones are too risky for rescue workers and first responders. Our aim with this solution is to provide a pre-insight to rescue workers before they are deployed to the actual zone. Camera-equipped drones and CCTV cameras can prove to be the primary sources of unstructured visual data from disaster-hit zones that can then be processed to extract meaningful insights which may help conduct search and rescue operations and also to prioritize disaster-hit zones based on the amount of damage occurred. As our model is very lightweight, we hope that it can be implemented on embedded systems and can be incorporated in pre-existing infrastructure with minimal changes.

For further testing, we have deployed our model on a local server and based on the images scraped from various online photo sharing platforms, we have tested our model's classification efficiency and inference time, and we have concluded that our model can successfully classify almost all the images into their respective classes and subclasses.

The proposed work has infinite scopes of development before it can be deployed in the real world. In future, we wish to extend our multi-label data set into more classes and finer subclasses and increase the number of images by scraping social media platforms as they contain images, taken on ground, during or after a natural disaster, which can prove to be very crucial for first responders. We will also be extending our research on using video streams because video streams provide much more richer information than static images.

## References

1. Mao J, Harris K, Chang N-R, Pennell C, Ren Y (2020) Train and deploy an image classifier for disaster response. In: Computer vision and pattern recognition
2. Chaudhuri N, Bose I (2019) Application of image analytics for disaster response in smart cities. In: 52<sup>nd</sup> Hawaii international conference on system sciences, pp 3036–3045
3. Rosebrock A (2017) Deep Learning for computer vision with python, PyImageSearch (2017)
4. Dogo EM, Afolabi OJ, Nwulu NI, Twala B, Aigbavboa CO (2018) A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In: International conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, pp 92–99
5. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) MobileNetV2: inverted residuals and linear bottlenecks. In: IEEE conference on Computer Vision and Pattern Recognition (CVPR)
6. Botev A, Lever G, Barber D (2017) Nesterov's accelerated gradient and momentum as approximations to regularised update descent. In: International joint conference on neural networks
7. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: International conference for learning representations, San Diego
8. Zhang Z (2018) IEEE/ACM 26<sup>th</sup> International Symposium on Quality of Service (IWQoS). Banff, AB, Canada, pp 1–2
9. Bock S, Goppold J, Weiß M (2018) An improvement of the convergence proof of the ADAM-Optimizer
10. Wang J, Cao Z (2017) Chinese text sentiment analysis using LSTM network based on L2 and Nadam. In: 17<sup>th</sup> IEEE International Conference on Communication Technology (ICCT), Chengdu, pp 1891–1895

11. Liu J, Strohschein D, Samsi S, Weinert A (2019) Large scale organization and inference of an imagery dataset for public safety. In: IEEE High Performance Extreme Computing Conference (HPEC), pp 1–6
12. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp 770–778
13. NOAA. Natural Hazard Images Database

# Developing a Portrait Drawing Robot with FANUC LR Mate 200id/4s



Ming Li

**Abstract** In this paper, a portrait drawing robot was developed based on FANUC LR Mate 200id/4s using Roboguide software. The goal was to enable the robot to draw a portrait simulating how a human being would draw, which is to draw a main shape of the portrait first and fill in details later. Fuzzy logic edge detection was utilized to determine the outline of a portrait. Membership functions and rules could be fine-tuned to adjust the intensity of detected edges so that the performance of the drawing robot could be improved. The drawing path of the robot was derived from the outlines identified by the edge detection and comprised a series of points. These points were targets of the robot arm, which could be visualized and validated through Roboguide. When the robot played back these teaching points, a portrait drawing could be completed.

**Keywords** Drawing robot • Fuzzy logic • Edge detection • FANUC robot • Roboguide

## 1 Introduction

Developing robots that can mimic human behavior on high-level tasks has been significant in attracting researchers. As artificial intelligence advances rapidly, it is possible that robots can integrate sensing, motion, and intelligence to accomplish what humans can do in their everyday lives. The state-of-the-art technologies enable robots to be utilized in a wide range of industries other than their traditional applications.

The entertainment industry has seen fast growth of interesting applications using autonomous robots. Researchers have devoted significant efforts to enhance creativity with technologies in order to endow robots with the ability to create high-quality artworks. A portrait drawing robot interacting with human beings and drawing their

---

M. Li (✉)  
University of Michigan-Flint, Flint, MI 48502, USA  
e-mail: [mingli@mich.edu](mailto:mingli@mich.edu)

portraits requires special human-like skills. Robots need to have the human level of perception, action, and cognition to achieve this challenging task.

A number of drawing robots have been developed during the last three decades [1, 2]. These robots can draw different types of paintings with the control of different end-effectors, including pencil sketch, watercolor painting, oil painting, etc. [3–5]. Some of these artistic robots can create human-like paintings which are very impressive. However, painting, as a way to express ideas and emotions as well as demonstrate aesthetic value, is abstract and complicated. Robots need to be well-trained to master the techniques and strategies that match with the level of human beings.

The aim of this paper is to develop a portrait drawing robot using a FANUC LR Mate 200id/4s. The robot mimics the human-like drawing style to produce a portrait. Through image processing, the robot was able to identify the main shape and details of the portrait drawing. Following the order of main shape first and details later, a sequence of teaching points was generated representing the path planning of the drawing. These teaching points needed to be validated to ensure that the robot was able to reach all of these poses. When the robot played back these teaching points with a pen connecting them, a portrait drawing could be completed.

The portrait to be drawn was processed in order to recognize the main shape and details of the drawing through functions of image processing. Edge detection was one of the main tasks in order to sketch the outline of the portrait drawing. In this paper, fuzzy logic edge detection was utilized to identify the outlines. Fuzzy logic processing is a powerful tool because it can implement rule-based expert systems using human intelligence. It is also robust since it does not need precise input, which is suitable to uncertainties and imprecision existing in many features of image processing [6, 7]. Fuzzy logic edge detection is desirable for portrait drawing also owing to its flexibility gained from fine-tuning the membership functions and rules.

The aforementioned fuzzy logic edge detection generated a series of points along the outline of the portrait to be drawn. These points were the poses for the End of Arm Tooling (EOAT), which is a marker held by the drawing robot. The path planning of the drawing robot was visualized and validated through Roboguide, which is a simulation software developed by FANUC [8, 9]. Roboguide supports offline simulation of operating a robot in a virtual work cell and provides features to create tool paths automatically [10]. In this case, the targets feature was utilized to represent and store points along the drawing path of the robot. Through simulation, the drawing path could be validated in order to confirm that the pen of the robot could move along each point in sequence to finish the drawing.

This paper is organized as follows: Sect. 2 describes the image processing and the fuzzy logic edge detection. Then, in Sect. 3, the targets feature in Roboguide is explained and the generation of the path of the EOAT is demonstrated. Finally, Sect. 4 concludes the paper with a brief summary.



## 2 Image Processing

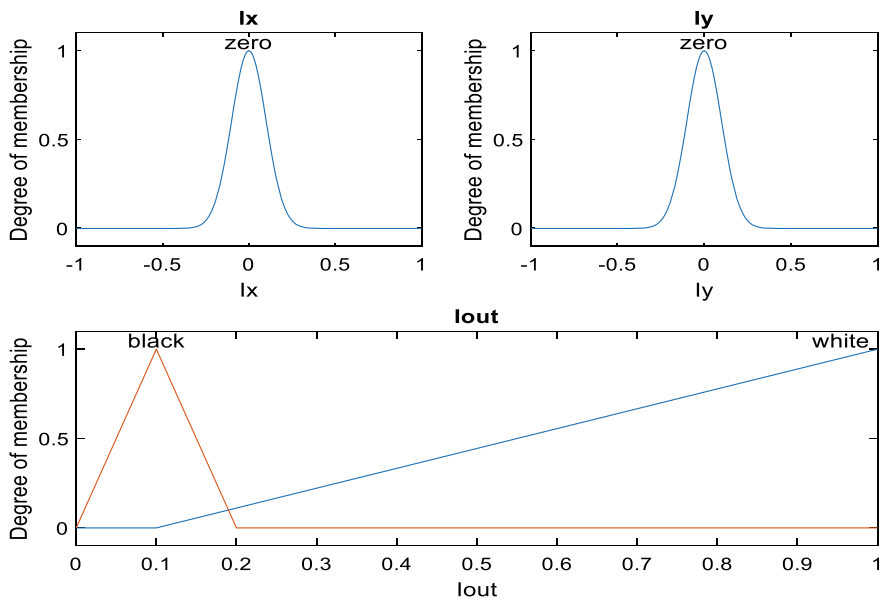
The image processing of a portrait to be drawn was based on functions of MATLAB Image Processing Toolbox. As shown in Fig. 1, an arbitrary picture of a portrait was chosen to be drawn by the robot. The fuzzy logic edge detection was utilized to identify the outline of the drawing. Figure 2 shows the membership functions of the fuzzy logic.  $I_X$  and  $I_Y$  are input membership functions along  $x$  and  $y$  directions and are Gaussian with a zero-mean. The standard deviation of the Gaussian can be changed to improve the performance of the edge detection. The membership functions for output are  $I_{out}$ , which determine if the output is white or black. They are triangular functions indicated by the values of start, peak, and end. The rules of the fuzzy logic identify the edges based on if the gradients along  $x$  and  $y$  directions are zero or not. If both gradients are zero then output is white, otherwise the output is black indicating the edge detected. Figure 3 shows the edge detection using different output membership functions. From the pictures, it is obvious that the membership functions of black are different, and the left one is much darker. The left membership function is  $[0 \ 0.5 \ 1]$ , and the right one is  $[0 \ 0.1 \ 0.2]$ . The right one with more white area was chosen, since the portrait was to be drawn by the robot.

After the edge detection, a binary picture was generated in which each pixel was either 1 or 0, indicating if that point was along the edge or not. The binary picture highlighted the outlines that the robot needed to draw. In order to draw the portrait in a human-like style, the robot followed the strategy of drawing main shape first and moving to details later. It identified the main shape by finding out longer chains of connected pixels and considered shorter chains of pixels as details. The robot scanned the binary picture until it found the first point which was 1, meaning it belonged to an edge segment. Then a recursive subroutine was called to detect a chain of pixels that connected together in all directions [11]. As shown in Fig. 4, the chains of pixels were the path planning for the pen in the robot arm to move along.

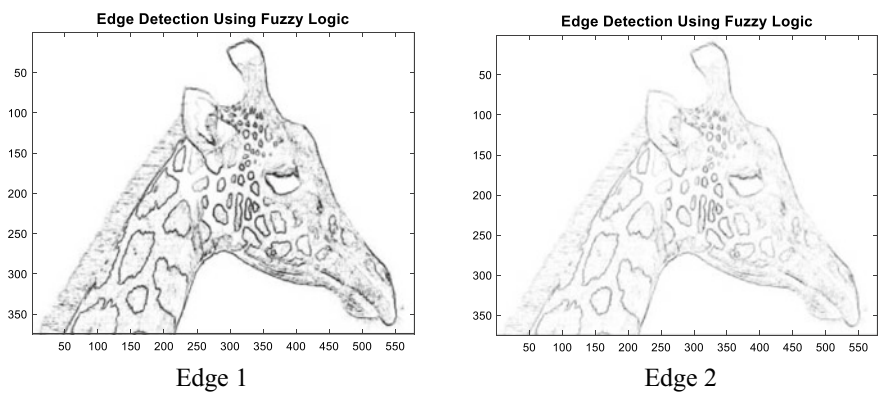
The FANUC LR Mate 200id/4s robot has six degrees of freedom and its workspace is defined by a Cartesian coordinate system attached to the base of the robot. The poses of an EOAT ( $x, y, z, w, p, r$ ) include the information of the position ( $x, y, z$ ) and the orientation ( $w, p, r$ ), which is the rotation angles of  $x, y$ , and  $z$  axes. Since the

**Fig. 1** Original picture





**Fig. 2** Fuzzy logic membership functions



**Fig. 3** Fuzzy logic edge detection

robot used a marker to draw, a tool frame was created to define the tip of the marker. Thus, the coordinates of the pixels identified from the image processing could be mapped to the position  $(x, y, z)$  of the EOAT. As for the values of orientation, a fixed set  $(0, 90, 0)$  was chosen because the pen needed to be perpendicular to the drawing area and the drawing was two dimensional. The robot aimed to draw like a human being and followed the order to draw the big shape first and fill in the details later. Through the image processing, the according chains of pixels had been laid out as a sequence of the EOAT poses of the robot.

### 3 Targets

The six degrees of freedom enable the robot EOAT to reach an arbitrary pose in a three-dimensional work space. However, given an arbitrary pose, it needs inverse kinematics to calculate the rotation degree of each joint to reach the desired pose. Occasionally, some poses cannot be reachable because they are singularities, out of axis limits, out of work space, etc. Validation of the working path is necessary to ensure that all the poses can be reached.

The targets feature of Roboguide can be utilized to visualize and validate the path planning of the robot. In the virtual working cell created in Roboguide, the targets represent the poses which are the desired points that the EOAT of the robot needs to move along. Depending on the total points along the working path of the EOAT, and the software running speed, it might take up to several minutes for Roboguide to convert the generated points into targets. Figure 5 shows the targets describing the poses for a portrait drawing to be accomplished. The green points illustrate the positions of the poses and the white triads attached to the green points show the orientation. However, among these green targets, there are red points as shown in Fig. 6. These red points are the poses that cannot be reached by the EOAT of the robot, and they must be removed or replaced. Through the adjustment and validation,

Fig. 4 Image processing

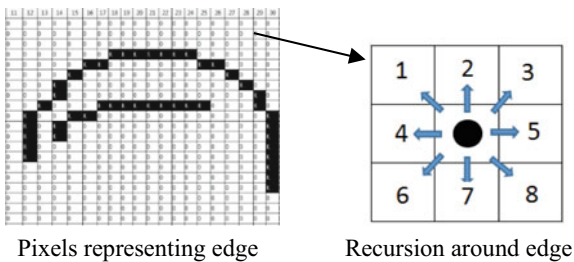
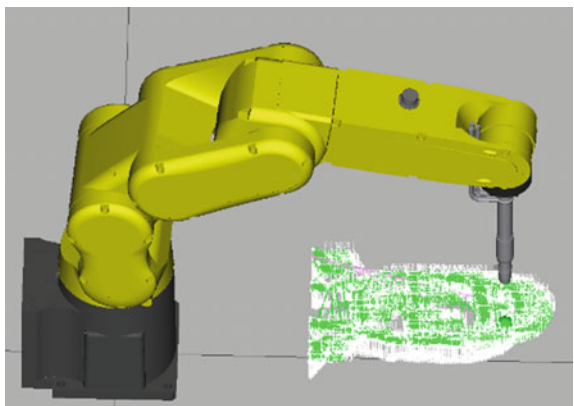
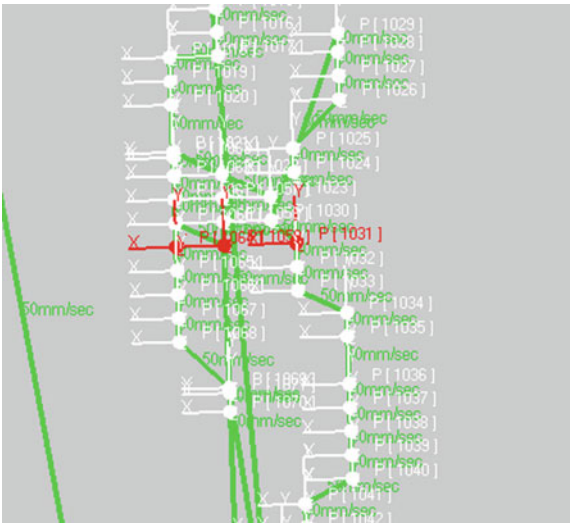


Fig. 5 Virtual work cell



**Fig. 6** Teaching points as targets



**Fig. 7** Portrait drawn by robot



an executable working path can be constructed. When the robot arm moved the pen along the desired poses, a portrait drawing could be accomplished. Figure 7 shows the drawing of an arbitrary portrait.

## 4 Conclusion

The project aimed to develop a drawing robot which could draw a portrait in a human like style based on a FANUC LR Mate 200id/4s. One of the main tasks of the robot was to identify the outlines of the portrait to be drawn. Fuzzy logic edge detection is powerful and flexible, so it was utilized to determine the outlines of the portrait. It is convenient to adjust membership functions and rules to achieve the desired

intensity of detected edges. The edges detected were transformed into a series of poses which constructed the drawing path of the robot. Then they were introduced into a virtual work cell created in Roboguide using the targets feature. The targets demonstrated the drawing path in the three-dimensional work space and validated that each pose could be reached in sequence. In this way, the robot played back poses of main shape first and details later to achieve the goal of drawing like a human. The proposed approach is flexible and fast and can be utilized in other applications. However, there are some limitations of this approach, e.g., some edge missed if the contrast was not strong enough, and the drawing was based on the pen perpendicular to the paper. In the future, other edge detection methods will be explored to improve the performance, and other drawing styles will be tested.

## References

1. Shih CL, Lin LC (2017) Trajectory planning and tracking control of a differential-drive mobile robot in a picture drawing application. *Robotics* 6(3):17–32
2. Gao F, Zhu J, Yu Z, Li P, Wang T (2020) Making robots draw a vivid portrait in two minutes. In: *IEEE/RSJ international conference on Intelligent Robots and Systems (IROS)*. IEEE Press, New York, pp 9585–9591 (2020)
3. Tresset P, Leymarie FF (2013) Portrait drawing by Paul the robot. *Comput Graph* 37(5):348–363
4. Scalera L, Seriani S, Gasparetto A, Gallina P (2019) Watercolour robotic painting: a novel automatic system for artistic rendering. *J Intell Rob Syst* 95(3):871–886
5. Beltramello A, Scalera L, Seriani S, Gallina P (2020) Artistic robotic painting using the palette knife technique. *Robotics* 9(1):15–35
6. Suryakant NK (2012) Edge detection using fuzzy logic in Matlab. *Int J Adv Res Comput Sci Softw Eng* 2(4):38–40
7. Kaur EK, Mutenja V, Gill IS (2010) Fuzzy logic based image edge detection algorithm in MATLAB. *Int J Comput Appl* 1(22):55–58
8. SimPro\_Master\_Help. ROBOGUIDE [computer software] help file. FANUC
9. ROBOGUIDE—FANUC Simulation Software. <https://www.fanucamerica.com/products/robots/robot-simulation-software-FANUC-ROBOGUIDE>. Last accessed 21 May 2021
10. Li L, Li X, Zhou X, Yue J (2007) Study of off-line programming system of arc robot based on the software of ROBOGUIDE. In: Tarn TJ, Chen SB, Zhou C (eds) *Robotic welding, intelligence and automation*, vol 362. Lecture notes in control and information sciences. Springer, Heidelberg, pp 401–408
11. Erik's XY-Plotter. [https://www.convict.lu/Jeunes/ultimate\\_stuff/Erik\\_s\\_xy\\_plotter/E\\_xy\\_plotter.htm](https://www.convict.lu/Jeunes/ultimate_stuff/Erik_s_xy_plotter/E_xy_plotter.htm). Last accessed 27 May 2021

# Comparative Analysis of Open Standards for Machine Learning Model Deployments



Sachchidanand Singh, Naveen Singh, and Vinay Singh

**Abstract** The process of adding a trained machine learning model into an existing production environment to make actual business decision is termed as machine learning model deployment. Deployment is entirely distinct from routine machine learning tasks like data pre-processing, feature engineering, model selection and model training, etc. The complete understanding and holistic view of how end-users are going to interact with the model's predictions is extremely important to decide how to deploy a machine learning model. But deployment process is rarely discussed and not very well understood amongst data scientists and machine learning engineers since they don't have DevOps background. This paper talks about different machine learning model types, their training, deployments and comparative analysis, insights into available open standards for machine learning model deployments to make informed decisions.

**Keywords** Machine learning (ML) · Open neural network exchange (ONNX) · Portable format for analytics (PFA) · Predictive model markup language (PMML) · State-action-reward-state-action (SARSA) · Deep Q neural network (DQN) · Q-learning · Hyperparameter · Return on investment (ROI) · Deep learning (DL) · Compound annual growth rate (CAGR)

## 1 Introduction

Machine learning algorithms help to discover patterns in data and provide guidance to build mathematical models using these findings. In machine learning world, the process of deploying models to production systems is often overlooked but it's one

---

S. Singh (✉)  
IBM India Software Labs (ISL), Pune 411045, India  
e-mail: [sachcsin@in.ibm.com](mailto:sachcsin@in.ibm.com)

N. Singh  
Tata Consultancy Services (TCS), Nagpur 440018, India

V. Singh  
Madan Mohan Malaviya University of Technology (MMMUT), Gorakhpur 273008, India

of the most critical aspect. There are very a few open and standard solutions available which provides guidance on the deployment of end-to-end ML pipelines.

The global machine learning (ML) market is expected to grow at a compound annual growth rate (CAGR) of 42.8% from 2018 to 2024 as per a Forbes report [1]. Adoption of Machine learning enabled solutions is growing rapidly due to return on investment (ROI) and better customer experience.

The scalability and deployments are tough challenges for AI models. Deploying ML models to production is a challenging task and takes a lot longer time. Organizations can become locked to a particular platform if a ML model lacks the capability to easily migrate to another host environment which in turn can create barriers for the data scientists [2].

Machine learning workflow which includes data cleaning, data preparation, model training, testing and deployments can be a long process with several issues along the way. Specially deploying and maintaining machine learning models at scale is critical to success of any data science project. Many data science projects fail to make it to production due to these challenges.

## **2 What is Machine Learning (ML) Model?**

Machine learning models takes request in the form of input data and make a prediction on the input data to serve a response. The machine learning models are trained over a set of data and algorithm suited for that kind of requirement to reason over and learn from the training data. Machine learning model are trained to recognize hidden patterns in underlying data.

A model is basically a mathematical representation of what a machine learning system has learned over time with training datasets and hence a model in machine learning is the output of ML algorithm run over data. After running a machine learning algorithm on the training data, a model is saved, and it represents the rules and data structures specific to the algorithm required to make the predictions.

The machine learning model contains model data and procedure for using the data to create a prediction, i.e., algorithm. Basically, algorithms are kind of optimization procedure to minimizes model errors on the training dataset.

## **3 Types of Machine Learning (ML) Models**

Machine learning can be categorized into supervised learning, unsupervised learning, or reinforcement learning. A supervised machine learning model can be further categorized either a regression or classification model.

The process of finding patterns and conclusions from input data without references to the labeled outcomes is called unsupervised learning. The two prominent methods used in unsupervised learning are dimensionality reduction and clustering. An agent learns in an interactive environment using feedback from its own actions and experiences in the reinforcement learning by trial and error method. For example, the most commonly used reinforcement learning (RL) algorithms are state-action-reward-state-action (SARSA) and Q-learning, etc.

### 3.1 *Supervised Learning*

In supervised learning, an algorithm is used to learn the mapping function from a given input variables ( $x$ ) to the output variable ( $Y$ ).

$$Y = f(x) \quad (1)$$

The training datasets contains inputs and correct outputs, which allow the model to train and learn over time. Supervised learning uses training data set to train models to produce desired output. It uses labeled datasets to train algorithms to predict outcomes or classify data accurately. The accuracy of algorithm is measured through loss function, and it is adjusted, fine-tuned so that error is minimized. Supervised learning can be broadly categorized into two types—classification and regression.

The process of identifying test datasets based on certain common features is called classification. It uses an algorithm to recognizes specific entities within the test dataset and accurately divide them into specific categories. Common classification algorithms are K-Nearest Neighbors (KNN), Decision Tree, Support Vector Machine (SVM), Naïve Bayes and Random Forest.

Regression is process of finding correlation between two sets of variables. It helps to predict the continuous output called dependent variable based on one or more predictor variables also called independent variables. It is commonly used for business forecasting of sales and revenue figures, weather forecasting like cyclone, rainfall, thunderstorm. The popular regression methods are linear regression, logistic regression, multivariate regression, and polynomial regression.

### 3.2 *Unsupervised Learning*

Unsupervised Learning allows machine learning model to work on its own to detect hidden information and data patterns. Unsupervised learning allows the model to work on its own without any supervision on unlabeled data. Unsupervised learning algorithms include k-means for clustering and Apriori algorithm for association rule.

Unsupervised learning can be broadly categorized into two types—Clustering and Association. The clustering is the process of dividing data points into several smaller



groups of similar characteristics so that data points in the same groups are more closely associated to each other. While association discovers patterns like customers that buy a given product A also tend to buy product B.

### ***3.3 Reinforcement Learning***

Reinforcement learning (RL) is used to find the best possible path or behavior in a specific situation. There is no correct output for reinforcement agent to decide best action for a given task and in the absence of a training dataset, agent is bound to learn from its own experiences. The reinforcement agent learns through trial and error to achieve a goal in a potentially complex environment and uncertain environment.

The learning agent receives positive feedback for each correct action and penalty or negative feedback for each wrong action. Here, agent learns automatically using feedback-based mechanism and primary goal of the agent is to improve the performance by acquiring maximum positive rewards.

Reinforcement learning are mainly two types—Positive Reinforcement and Negative Reinforcement. Reinforcement learning algorithms are mainly used in gaming applications. The process of adding something to increase the tendency that expected behavior would reoccur is called positive reinforcement learning. It increases the strength of the behavior. Negative reinforcement learning increases the tendency that a specific behavior will reoccur by avoiding the negative condition. The most used algorithms are deep Q neural network (DQN), Q-Learning and State-Action-Reward-State-action (SARSA).

## **4 Machine Learning (ML) Model Training and Prediction**

Machine learning (ML) training involves training the model with training datasets and machine learning algorithm. The training dataset contains correct answer called as target attribute and learning algorithm finds specific patterns in the training dataset to map input data to the correct target attributes.

In training process, random values are initialized for each of the training matrixes and then attempts to predict the output of the input data using the initial random values. Weight and bias are learnable parameters, weight decide how much influence a given input will have on the output, whereas bias delays the triggering of the activation function by shifting to the left or right, to better fit the data. The weights and biases values of ML model is adjusted until model starts giving good predicting results. This process is repeated, and, in each iteration, the initial random line moves closer to the more accurate results.

## ***4.1 Splitting into Training and Testing Datasets***

The available dataset is split into training and test data. The training dataset contains a known output and the model learns on this training data. The prediction and accuracy of machine learning model is validated on testing dataset.

The train-test data split having higher proportion of data in the training set is more likely give better accuracy as calculated on that test dataset. You can make it 80:20, 75:25, 60:40 for getting better results.

## ***4.2 Splitting into Training, Test and Validation Datasets***

In this approach 50% or higher percentage of the data is allocated to the training set, 25% to the test set, and the remaining portion to the validation set. Thumb rule is to never train ML models on test data. If you are getting extremely good results, then probably you trained your ML model on the test set or test data has leaked into your training set.

A machine learning algorithm takes two types of input—model parameter and model hyperparameter. A model parameter is a configuration variable which is internal to the model. It's basically weight of the connection in the context of neural network and during training stage these parameters are learned and tuned.

While model hyperparameter is a configuration variable which is external to the model and its value can't be estimated from data. The hyperparameters for a neural network are typically number of the hidden layers, learning rate, number of epochs, etc. Grid Search is used to find the best combination of hyperparameters to yield the most accurate predictions.

### **Training Phase**

- Transform Data
- Extract Features
- Pre-process Features
- Train Model.

### **Predication Phase**

- Transform Data
- Extract Features
- Pre-process Features
- Prediction.

## 5 Open Standards for Machine Learning (ML) Model Deployments

The method of adding or integrating your newly developed model into the existing production environment is called ML model deployment. Developing a model and using the model in a real-time or production environment is never an easy task as it requires a lot of factors to be considered. A good model should be portable and scalable. Scalability of the model defined by the ability to distribute the load and computation based on input data size. Portability of the model is defined by the ability to execute in another environment with minimum interference. Portability and scalability both dependent upon the underlying language and hardware heavily [3, 4].

### 5.1 *Predictive Model Markup Language (PMML)*

Predictive model markup language (PMML) is an open standard for XML representation. It was developed by the Data Mining Group (DMG), a council of open-source data mining companies. PMML allows easy movement or migration of predictive solutions to the different tools and applications without the need of custom coding [5–7].

PMML allows developers to develop a model in one system and deploy using another application. PMML mainly contains—Header, Data Dictionary, Transformation Dictionary and Models. PMML is widely used for model visualization, scoring, and easily exchange models between products, etc. PMML supports a huge range of models and languages.

PMML is basically an XML-based standard used to represent statistical and data mining predictive models. A PMML file may contain a set of business rules to define decisions based on predictive model outcome. PMML has several built-in functions which allow for extensive data manipulation like arithmetic functions and IF–THEN–ELSE. It defines an element to represent multiple models like model segmentation, cascading and composition, etc. All the capabilities available for data pre-processing are available for post-processing with PMML 4.1.

PMML model contains following information about the model:

**Mining Schema**—It includes information about the predictors and target attributes.

**Model information**—It is used for model interpretation and visualization.

**Local Transformations**—It holds information about transformation in the model.

**Model verification**—It is used to store the information for correctness of the model and parameter matrix.

**Output**—scoring outputs and post-processing.

#### Benefits of PMML

- Human readable
- Support for many models
- Support wide range of models
- Easy to use.

### **Issues**

- XML format is obsolete, and industry is moving away from it.
- Does not support deep learning.
- Extensions are not portable.

## **5.2 *Portable Format for Analytics (PFA)***

Portable Format for Analytics (PFA) is a framework which helps in smooth transition between multiple environments. PFA document is a JSON-based serialization of a scoring engine or a model. PFA does not change the underlying environments but acts as a bridge between two environments [8–10].

Portable Format for Analytics (PFA) was developed by Data Mining Group (DMG), an association managed by the Center for Computational Science Research, Inc. PFA facilitates easy portability across systems with algorithmic flexibility and an emerging standard for data transformation engines and statistical models. The pre-processing and post-processing functions can be randomly composed and built into workflows using PFA.

PFA is a JSON-based predictive model interchange format having comprehensive built-in function library. PFA have built-in support for some common models like clustering, linear models, and decision tree.

### **Benefits of PFA**

- User defined functions
- Flexibility
- Strong support for traditional ML operations.

### **Issues**

- No built-in support for mixed dense/sparse vectors
- No built-in support for generic tensors (3D+)
- No built-in function for typical deep learning models (e.g., CNN, RNN)
- No support/awareness of GPU.

## **5.3 *Open Neural Network Exchange (ONNX)***

Open neural network exchange (ONNX) is the newest open standard. It was first released by Microsoft and Facebook for deep learning models in 2017, later augmented to support “traditional” machine learning. ONNX is an open format that

makes machine learning models more portable. ONNX has joined Linux Foundation AI [11–14].

It is currently being actively developed by several companies and individual contributors. In this talk, we explore the differences of ONNX from earlier open standards PMML and PFA. We will explain their different approaches to data preparation and model representation, compare their features, and point out opportunities for open-source contributions.

ONNX framework allows interoperability of various models which could be written or trained using different frameworks. Each machine learning framework is tuned to work best for a certain part of the model building process. For example, Spark ML supports distributed computing framework and works best in the case of large dataset, whereas Caffe2 is known for speed in the case of deep learning computing.

If a user wants to develop a model using Keras and deploy using PyTorch then such cases can be easily handled using the ONNX framework. If we want to support ONNX-ML, we need a PMML to ONNX converter, after adding all missing parts into ONNX.

## 6 Summary and Conclusion

Organizations are significantly adopting machine learning enabled solutions to enhance customer experience, competitive edge and return on the investment. The use of machine learning applications in many industry verticals is likely to grow exponentially in the coming years fueled by enormous data generation and technological advancements. As per report from a market research company MarketsandMarkets, the machine learning market is anticipated to grow at a compound annual growth rate (CAGR) of 44.1% from USD 1.03 Billion in 2016 to USD 8.81 Billion by 2022 [15].

With rapid adoption of machine learning in industry, there is huge demand for language-agnostic exchange formats, open-source driven tools, open format for ML models and framework independent file formats. Some examples of open standards are predictive model markup language (PMML), open neural network exchange (ONNX) and portable format for analytics (PFA). Facebook and Microsoft developed open neural network exchange (ONNX) to serve as open standard for machine learning model interoperability. It's used for exchanging deep learning (DL) models between different deep learning frameworks.

The data science projects have historically high failure rates. As per VentureBeat report, 87% of data science projects are never deployed on production [16] and a Gartner report suggest that only 15 to 20% of data science projects get completed [17, 18]. But to succeed in the fast-changing business world, there must be a holistic understanding on architectural view and how data science projects will be deployed, integrated within the production pipeline, and deliver business value.

## References

1. Roundup of machine learning forecasts and market estimates (2020). <https://www.forbes.com/sites/louiscolombus/2020/01/19/roundup-of-machine-learning-forecasts-and-market-estimates-2020/?sh=14bb20675c02>
2. Challenges in deploying machine learning: a survey of case studies. <https://arxiv.org/abs/2011.09926>
3. Open standards for machine learning model deployment. [http://www.data-science-nights.org/uploads/8/4/8/7/84875762/deployml\\_sep\\_25.pdf](http://www.data-science-nights.org/uploads/8/4/8/7/84875762/deployml_sep_25.pdf)
4. Overview of the different approaches to putting Machine Learning (ML) models in production. <https://medium.com/analytics-and-data/overview-of-the-different-approaches-to-putting-machinelearning-ml-models-in-production-c699b34abf86>
5. Open standards for deployment, storage and sharing of predictive models. PMML/PFA/ONNX in action. [https://www.researchgate.net/publication/334611859\\_Open\\_standards\\_for\\_deployment\\_storage\\_and\\_sharing\\_of\\_predictive\\_models\\_PMML\\_PFA\\_ONNX\\_in\\_action](https://www.researchgate.net/publication/334611859_Open_standards_for_deployment_storage_and_sharing_of_predictive_models_PMML_PFA_ONNX_in_action)
6. Productionizing spark ML pipelines with the portable format for analytics. <https://databricks.com/session/productionizing-spark-ml-pipelines-with-the-portable-format-for-analytics>
7. Facebook, microsoft bring interoperable models to machine learning toolkits. <https://thenewstack.io/facebook-microsoft-bring-interoperable-models-machine-learning-toolkits/>
8. Deploying your machine learning model to unlock its potential. <https://s3.us.cloud-object-storage.appdomain.cloud/developer/videos/deploying-your-machine-learning-model-to-unlock-its-potential-tech-talk/static/deploying-ml.pdf>
9. The evolution from PMML and PFA to agnostic scoring engines. <https://www.modelop.com/blog/the-evolution-from-pfa-and-pmml-to-agnostic-scoring-engines/>
10. Productionizing ML pipelines with PFA. <https://www.bigthingsconference.com/2018/talk/Productionizing-ML-Pipelines-PFA/>
11. 5 Machine learning trends for 2018 combined with Apache Kafka ecosystem. <https://dzone.com/articles/5-machine-learning-trends-2018-combined-with-apache>
12. Deploying analytics with the Portable Format for Analytics (PFA). <https://doi.org/10.1145/2939672.2939731>
13. Deploying machine learning models for edge computing on drones. <https://blog.dataiku.com/deploying-machine-learning-models-for-edge-computing-on-drones>
14. Deploying machine learning models in practice. <https://qconsp.com/sp2018/system/files/presentation-slides/qconsp18-deployingml-may18-npentreath.pdf>
15. Machine learning market by vertical (BFSI, healthcare and life sciences, retail, telecommunication, government and defense, manufacturing, energy and utilities), deployment mode, service, organization size, and region—global forecast to 2022. <https://www.marketsandmarkets.com/Market-Reports/machine-learning-market-263397704.html>
16. Why do 87% of data science projects never make it into production? <https://venturebeat.com/2019/07/19/why-do-87-of-data-science-projects-never-make-it-into-production/>
17. Our top data and analytics predicts for 2019. [https://blogs.gartner.com/andrew\\_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/](https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/)
18. 10 Reasons why data science projects fail. <https://fastdatascience.com/why-do-data-science-projects-fail>

# Role of ICT in Emerging Towns and Cities



S. Deepthi , R. Shanthi Priya, and S. RadhaKrishnan

**Abstract** Urbanization at unprecedented growth demands innovative and proactive response for better place to live. Information communication and technology have changed the concept of spatial planning in towns and cities. Successful application have been effectively used in cities and towns developed by private and public agencies for improving the living environment. A wide range of practices are discussed in this paper based on the themes of urban infrastructure, public services, urban governance, energy and environment, public health and safety practiced at national and international level. The ICT tools and urban parameters used in case areas would acknowledge the framework for the successful applications. Considerably, ICT services with governmental mechanism and policies can act as a facilitator for the success of the project. However, real-time information plays an important role for data collection which supports effective service to the public.

**Keywords** Spatial planning · ICT tools · Urban infrastructure · Urban governance · Public health and safety · Energy and environment

## 1 Introduction

Cities and towns have become more complex due to increase in population and high demand on infrastructures, leading to enormous strain on urban fabric. Urban infrastructure including utilities and transportation is under maintained. The sectors of social services, education, and healthcare are difficult to sustain because of the problem associated. Traditional methods have become outdated to tackle issues at a large pace. Model with technologies are advancing in rapid pace. A new intelligent management is much needed for urban and spatial planning. Sustainable development

---

S. Deepthi (✉) · S. RadhaKrishnan  
Prime Nest College of Architecture and Planning, Trichy, Affiliated To Anna University, Chennai, Tamilnadu, India

R. Shanthi Priya  
School of Architecture and Interior Design, SRM Institute of Science and Technology, Kattankulanthur, Tamilnadu, India  
e-mail: [shanthir1@srmist.edu.in](mailto:shanthir1@srmist.edu.in)

goals strongly highlight the need for accessing information and communication to build resilient infrastructure, to promote strong inclusiveness, to enhance sustainable industrialization, and to develop innovation [1]. PWC report on sustainable urbanization states that world needs a new operating system which provides residents a sustainable way of economic viability, social inclusion, and environmentally sustainable way of living [2]. City investment includes large ICT which enables designing of smarter cities for better quality of life. Major cities across the world have started using ICT for providing services and good quality of life to citizens. Urban environment is in close linkage with technologies. ICT in spatial planning can be conceptualized as a new approach for transferring of data and information. The new concepts evolve when technology creates impact on urban environment. The concept of ICT has been introduced through various concepts like smart cities, intelligent cities, and digital cities as a solution for current urbanization phenomena [3]. Implementing ICT in cities and towns would create opportunities to manage real-time scenarios like transparency in governance, infrastructure management, increasing the efficient use of resource, and mitigating climatic changes [4]. Currently, urban planners are utilizing technology around the globe in various aspects to monitor day-to-day needs while encouraging the capacity to forecast and acknowledge the urban issues.

## **2 Role of ICT in Towns and Cities**

See Fig. 1.

### ***2.1 Real Estate and Building***

The role of smart cities in real estate industries benefits the city community while selecting the location of the project, ensuring city network for public participation for future building perception as well as creating open source data provided by the government during design developmental stage [5]. Integration of building modeling system and geographical information system plays a major role in enabling three-dimensional city model with data offered from Internet of things (IOT), big data centers, and sensors established by smart city governing system [6]. The framework for evaluating real estate development based on smart building creation provides greater possibilities to assess the existing condition as well as predict the future real estate development [7].



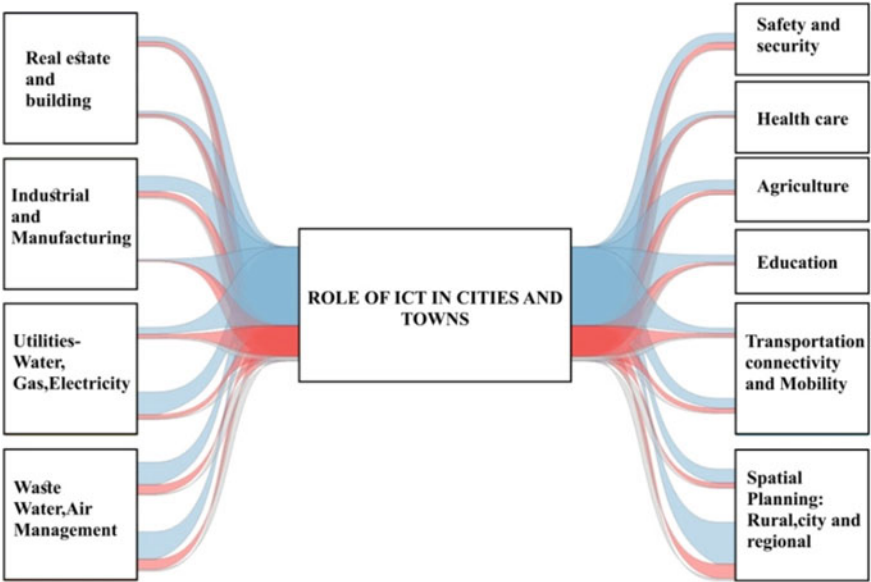


Fig. 1 Role of ICT in various sectors of cities and towns

## 2.2 Industries and Manufacturing

The application of ICT in industries and manufacturing sectors is in the early stage of development [8]. [9] Recommends three dimensions for evaluating ICT practices: operational leadership, marketing efficiency and relationship between them with ICT for reliable measurement of ICT's efficiency in an organization.

## 2.3 Utilities—Water, Gas, Electricity

According to World Energy Council, ICT has played a major role in managing supply and demand of energy monitored by smart metering at household level. Based on case study conducted in Midland regions of UK, the reduction analysis presents home energy management System and ICT could potentially reduce energy consumption by 33% [10].

## 2.4 Waste Management

In order to maintain a healthy community living, importance of waste management, recycling, and modern equipment for collecting and processing must be implemented

by the government to reduce waste impact on environment. ICT technologies for waste management becomes mandatory because of the rigorous problem associated with waste generation at cities and towns. The concept of cloud computing, remote sensing, GIS, virtualization, online Web service and technologies based on computer-aided design, computer-aided engineering, and computer-aided manufacturing have been implemented in France for handling waste generated [11].

## ***2.5 Safety and Security***

Public safety and security has always been a concern in day-to-day life. Analytics and surveillance to monitor can ensure public safety and prevent crime. As per JLL report on smart cities' success report: smart surveillance, predictive policing, mapping of crime, resource optimization, disaster altering, and detecting systems are few measures to reduce crime. Real Time Crime Centre is technology-based solutions that are world-wide successful. By analytical and visualization tool New York police department Real Time Crime Centre has reduced crime rate by 27% as per PWC report.

## ***2.6 Health Care***

Tele health facilitates helps to connect hospitals to remote areas for the purpose of consultation, training, and diagnosis to increase accessibility. Southwest china has implemented tele-medicine solutions in two hospitals in Shapingba district which is functioning at a greater extent.

## ***2.7 Agriculture***

Efficient management of supply chain with real-time information would enhance the performance of the entire supply chain process. Introduction to ITC-Choupal in India is the most successful model equipped with PC, Internet connectivity with Vsats, power supply with solar energy, and printer [12]. ITC-Choupal strongly established rural hubs and created complete information system for farming communities to increase productivity.

## **2.8 Education**

The rapid change has been witnessed in the field of education due to information communication and technology. Smart schools and universities have created platform for virtual learning. Currently, online learning and training program has become the new norm of teaching [13]. The pattern of IT education has evolved as a platform of knowledge sharing, techniques, and controlling standards in the society.

## **2.9 Mobility**

To ensure integrated infrastructure to improve efficiency high quality mobility is needed for the success. Mobility is used as an interactable and a major challenge faced by many cities across the globe. Intelligent transport system comprises of different kinds of operational technologies used for managing transport, sensors, online ticketing infrastructure, and surveillance which are employed along with the travel route to manage and monitor travel behavior. Crowdsourcing techniques like Google Map, TomTom's map share, NAVTEQ captures real-time travel data for the use of functioning control [14].

### **2.9.1 City Planning**

The influence of information communication and technology in spatial planning is noteworthy. Introduction to digitized mapping technique through satellite images, large-scale 3d models, risks of urban floods, policy making, transportation planning, and natural resource management are developed in the context of ICT with wide large of information [15].

## **3 International Case Studies on ICT-Spatial Planning Context**

See Table 1.

## **4 National Case Studies on ICT-Spatial Planning Context**

See Table 2.

**Table 1** Role of ICT in different sectors of spatial planning in International context

Public services and government	<ul style="list-style-type: none"> <li>• Case area 1: Beijing, China</li> <li>- Agency: Dongcheng district Government</li> <li>- Urban Parameters: Street lighting, drainage, water supply, underground pipelines, housing, gardens, and manholes covers</li> <li>- ICT tools: GIS, GPS technology by gridding methods, and mobile technology</li> <li>Outcomes: Tracking of governmental services relation between municipal administration and citizens</li> <li>• Case area 2: Fix my street, London</li> <li>- Agency: My society not a profit company</li> <li>- Urban parameters: Street lights, street, and roads</li> <li>- ICT tools: Fix my street app developed in 2008</li> <li>Outcomes: E-participation increased</li> </ul>
Urban administration	<ul style="list-style-type: none"> <li>• Case area: Institution knowledge and Technology Management, Peru</li> <li>- Agency: Regional government of Callao</li> <li>- Urban parameters: Information on historic monuments, disaster risk management</li> <li>- ICT tools: ICT-GIS-based knowledge management system for local and regional development</li> <li>Outcomes: Supporting of spatial planning and management with updated data on spatial information</li> </ul>
Public health and safety	<ul style="list-style-type: none"> <li>• Case area: Kenya Zambia</li> <li>- Agency: Family health International of Sub-sharan Africa</li> <li>- Urban parameters: Predict the spread of disease and identify health gap</li> <li>- ICT tools: GIS with GPS data resources</li> <li>Outcomes: Resource allocation was possible through GIS which facilitated policy making associated with health care</li> </ul>
Environment and energy	<ul style="list-style-type: none"> <li>• Case area: Germany, Solar Potential Mapping</li> <li>- Agency: German Association for surveying and Mapping</li> <li>- Urban parameters: Roof top data</li> <li>- ICT tools: Aerial laser scanner, ArcGIS spatial analyst, ArcGIS desktop model builder</li> <li>Outcomes: Solar energy installation increased. GIS can be used for creating renewable energy resource map</li> </ul>

(continued)

**Table 1** (continued)

Urban infrastructure planning	<ul style="list-style-type: none"><li>• Case area 1: Singapore<ul style="list-style-type: none"><li>- Agency: Singapore Land Transport Authority</li><li>- Urban Parameters: Public transport road, road operation and road network</li><li>- ICT tools: Electronic road pricing, symphony for e-payment, parking guidance system</li><li>Outcomes: Shorter delay of travels and lower traffic congestion</li></ul></li><li>• Case area 2: Real-time water supply management in Bangkok<ul style="list-style-type: none"><li>- Agency: Thailand’s metropolitan waterworks</li><li>- Urban parameters: Water distribution, water treatment plant</li><li>- ICT tools: Magnetic and ultrasonic flow meter with provision of field control junction, intelligent remote terminal units with GPRS support system</li><li>Outcomes: Reliability and accuracy of water distribution system helps to control and reduce losses</li></ul></li></ul>
-------------------------------	---

*Source* Learning’s from PERAL, Compendium of Global Good Practices, ICT in Urban Services

**5 Discussion**

The application of information technology and communication in various sectors of cities and towns has helped in fostering livable communities and creating better quality of life. The practices of ICT in many cities nationally and internationally would have encountered challenges before success of the project. Both nationally and internationally ICT has evolved concepts and solutions for infrastructure services, Administration, business, and governance to increase problem solving ability of the city. Mostly, the successful innovation of technologies in International cities are outcomes of smart cities information system (SCIS) involving developers, planners, software analyst, institution, and communities across the country [16]. In India, the concept of modernization was initiated through JNNURM scheme under the Ministry of Urban Development, 2007. Later, smart cities mission which was launched in 2015 promoted infrastructure and sustainable development through smart solution [17]. Based on six concepts of smart cities, integration of innovative methods has been implemented in sectors of solid waste management and water management, transportation, governance, education, agriculture, and health and safety. Further, the concept of low cost ICT applications have greater potential for future development.

**6 Conclusion**

The paper discusses about the real-time use of ICT in different sectors of spatial planning at national and international context. The success implementation will help us to understand the context base study, collecting real-time data, creating awareness, and

**Table 2** Role of ICT in different sectors of spatial planning in national context

Environment and energy	<ul style="list-style-type: none"> <li>• Case area: Flood management cyclone</li> <li>- Urban parameters: Location-based information of existing infrastructure information on roads, landsite area, prediction of severity and uncertainty</li> <li>- ICT tools: GIS-based mapping with real-time satellite maps, VHF radio communication for fishermen community in Tamilnadu, Relief management App</li> </ul> <p>Outcome: Practices help for preparedness at local and national level</p>
Urban infrastructure	<ul style="list-style-type: none"> <li>• Solid Waste Management</li> <li>• Case area: Jabalpur, Madhya Pradesh and Indore</li> <li>- Urban parameter: Garbage collection, transportation, and processing disposal waste</li> <li>- ICT tools: Creation of dashboard, ITC tools tracking monitoring, RFID semi underground bins through bin level sensor, and vehicle tracking and monitoring system (VTMS)</li> </ul> <p>Outcome: Effective control and management of disposal of solid waste at household level</p> <ul style="list-style-type: none"> <li>• Transportation</li> <li>• Case area: Ahmedabad, Janmarg BRT</li> <li>- Urban parameter: Socio-economic characteristics, travel demand patterns, road network, and proposed metro plan</li> <li>- ICT tools: GPS-enabled buses control center for traffic management, passengers information system</li> </ul> <p>Outcome: Continuous effort is made to achieve sustainable transport with pedestrian bicycle lane</p> <ul style="list-style-type: none"> <li>• Water supply</li> <li>• Case area: Area-based proposal for <math>24 \times 7</math> water supply for Raipur city under Smart City Mission</li> <li>- Urban parameter: Operation and maintenance of pipeline network considered area of study, identification of non-revenue connection expert service for leak detection, and operation and maintenance with metered connection</li> <li>- ICT tools: Mapping of water supply zones as GSM/GPRS technology for monitoring SCADA (Supervisory Control And Data Acquisition)</li> </ul> <p>Outcome: Development and Implementation in process</p>
Urban governance	<ul style="list-style-type: none"> <li>• Case area 1:Jabalpur M-governance as per Smart city proposal</li> <li>- Urban parameters: Birth and death certificate, payment of property tax, access to service such as water supply approval of building plan, women and senior helpline</li> <li>- ICT tools: Jabalpur 311 mobile-based app for smartphones</li> </ul> <p>Outcome: Quickened the easy and service to civic facilities among citizens</p> <ul style="list-style-type: none"> <li>• Case area 2: Bhubaneshwar One-Bhubaneshwar</li> <li>- Urban parameters: locations, ward information, public service information of tourist site etc.,</li> <li>- ICT tools: GIS-based map with integration of geo-spatial data</li> </ul> <p>Outcome: One stop solution to residents of the city</p>

(continued)

**Table 2** (continued)

Urban health	<ul style="list-style-type: none"> <li>• Case area 1: Public Bicycle-Bhopal               <ul style="list-style-type: none"> <li>- Urban parameters: Bicycle track connecting public transit nodes covering 12 km stretch, 50 locations were identified for hiring bicycle</li> <li>- ICT tools: GIS-based map</li> </ul> </li> <li>Outcome: 50% of the users were registered currently</li> <li>• Case area 2: E-Toilets, Solapur               <ul style="list-style-type: none"> <li>- Urban parameters: Based on floating population, public convenience</li> <li>- ICT tools: Display and instructions boards, automatic flush and cleaning mechanism with real-time electronic monitoring</li> </ul> </li> <li>Outcome: Cleanliness and hygiene of the city. Open urination has been avoided</li> </ul>
--------------	---

*Source* Learnings from SMARTNET, Ministry of Housing and Urban Affairs, Government of India

training program on ICT-enabled services, establishing management system. Furthermore, the paper describes the urban parameters which has enabled for real-time data collection in each urban sectors for better implementation and service optimization. Challenges like operational efficiency, creating systemic capacity, investment, constant monitoring of real-time information etc., require more attention. Though implementation and use of ICT-enabled tools track records of successful projects more aspects like communication infrastructure, transparency in governance, efficient use of resources, service delivery, user friendly etc., have to be enhanced to deal with long standing urban issues. In India, several cities have deployed the use of smart techniques through schemes and policies. However, we have long way to go in establishing the digitalized network. The future work of ICT in India can be extended to solve several problems in mitigating climatic changes, predict the risk of spread of diseases, hazard risk mapping, usage of renewable resources etc., through technologies like remote sensing, developing spatial database and e-government infrastructure.

## References

1. Whitaker F Sustainable development goal 9: investing in ICT access and quality education to promote lasting peace. <https://www.un.org/sustainabledevelopment/blog/2017/2021/04/24>
2. PWC (2010) Strategies and sustainable urbanization and role of ICT in city development. Mumbai
3. NIUA (2015) ICT in urban service. New Delhi
4. (2020) Smart cities and inclusive growth. France
5. Kumar S, Talasila V, Pasumarthy R (2021) A novel architecture to identify locations for Real Estate Investment. Int J Inform Manage 102012
6. Panteli C, Kyli A, Fokaides PA (2020) Building information modelling applications in smart buildings: from design to commissioning and beyond a critical review. J Cleaner Prod 121766
7. Apanaviciene R, Urbonas R, Fokaides PA (2020) Smart building integration into a smart city: comparative study of real estate development. Sustainability 12(22):9376

8. Tewari SK, Misra M (2012) The impact of ICT on manufacturing industry: an empirical analysis. In: International conference on communication systems and network technologies, Rajkot, Gujarat. IEEE, pp 924–929
9. Aldhmour F, Shannak R (2009) The effective utilization of information and communication technology and its impact on competitive advantage. 29(3):302–314
10. Cosar-Jorda P, Buswell R, Mitchell V (2015) Identifying the opportunities for ICT based energy demand reduction in family homes EEDAL: energy efficiency in domestic appliances and lighting. In: Lough borough University. European Commission under the Creative Commons Attribution, Lucerne-Horw, Switzerland
11. Ion I, Gheorghe FF (2014) The innovator role of technologies in waste management towards the sustainable development. *Procedia Econ Finance* 8:420–428
12. Ali J, Kumar S (2011) Information and communication technologies (ICTs) and farmers decision-making across the agricultural supply chain. *Int J Inform Manage* 31(2):149–159
13. Alfahad FN (2012) Effectiveness of using information technology in higher education in Saudi Arabia. *Procedia-Soc Behav Sci* 46:1268–1278
14. Aoun C et al (2014) Urban mobility in the smart city age
15. Wang TAO (2013) Interdisciplinary urban GIS for smart cities: advancements and opportunities. *Geospatial Inf Sci* 25–34
16. (2017) The making of a smart city: best practices across Europe. Belgium
17. Smart Cities. <http://smartcities.gov.in/,2021/04/24>



# A Web-Based Application for Data Collection and Report Generation Using Django



C. Arun, S. Priyenka, L. Sagarika, and S. Sneha

**Abstract** This paper proposes an user friendly Graphical User Interface (GUI) for data collection, maintaining databases and automated report generation processes. In recent years, the use of database has become more prevalent in almost every industry. However, the lack of technical knowledge in handling databases is common among most of the end users. More importantly, key data collection is a critical task while handling huge database research. This system makes the data collection process very simple by enabling the authorized users to update the database with Django Authentication System and also reduces the administrator's load of frequently updating the database. Data generation includes tasks such as searching, focusing, monitoring, selecting, extracting, and capturing data. The proposed reporting software is more sophisticated and efficient in the reporting process than spreadsheets, from data entry to display. This application uses HTML and CSS in the user interface, JavaScript for interactivity, Microsoft's MySQL for backend and Django web-framework for processing. With the combination of the above mentioned tools, the proposed system is successful in collecting, storing, manipulating and generating customized reports based on the user's request. It is intelligent enough and has flexible methods to fetch data using dynamic SQL queries in generating reports with multiple parameters.

**Keywords** Database · Extracting · Data collection · Reports

## 1 Introduction

Database is gathering of facts and statistical information collected and stored for analysis. Database is the key in today's world across all major commercial establishments and services. Like, Government sectors, e-commerce, banking, Universities, health care etc. Data collection is the process of collecting, measuring and analyzing accurate research data using standardized methods [1]. In many cases, data collection

---

C. Arun (✉) · S. Priyenka · L. Sagarika · S. Sneha

Department of Electronics and Communication Engineering, RMK College of Engineering and Technology, Chennai, India

is the first and most important step, regardless of the field of study. The data collection method varies from different learning areas, depending on the details required [2]. The most important purpose of data collection is to ensure that rich and reliable data is collected for statistical analysis so that data-driven decisions can be made for research. Report production is an important part of developing a business application. This application comprises of advanced computing system where reports with various combinations of key parameters are generated. It is built with SQL reporting services which is one of the most popular and proven reporting tools. The applicative solutions that we propose enables a common user to update data and also create own customized report and gathering of information.

This application with user friendly interface will help both the advanced and novice user comfortable to work with. This web based application with simple GUI, does not expect the user to have prior technical knowledge in DBMS or database query language which the novice user may not be well versed with [3].

The others are the more advanced users, more experienced database administrators with high knowledge of DBMS. For them it serves as a faster and more convenient method in updating the databases frequently, and generating a customized report effortlessly.

## **2 Implementation**

### **2.1 Web Application**

This concept is implemented as a web based application and it enables the users accessibility online 24/7. The implementation is done in such a way that it reduces load on the administrator's side in maintaining the updates made by the users [4].

### **2.2 Development Tools**

The database tool used in the application is Microsoft's MySQL8.0 which is connected to the web application using a Python-based free and open-source web framework that follows the model-template-views architectural pattern called as Django. HTML (Hypertext Mark-up Language) and CSS (Cascade Styling Sheet) for developing the front-end and the scripting language used for interactivity with the users is JavaScript [5, 6].

### 3 Workflow

The workflow of the proposed system is shown Fig. 1. This system preferably can be used in educational institutions to automatically update the publications of the professors into the database and also helps in generating a customized report of the same whenever required.

The application has 4 major tabs in the navigation bar after logging in with the appropriate credentials.

#### 3.1 Login Page

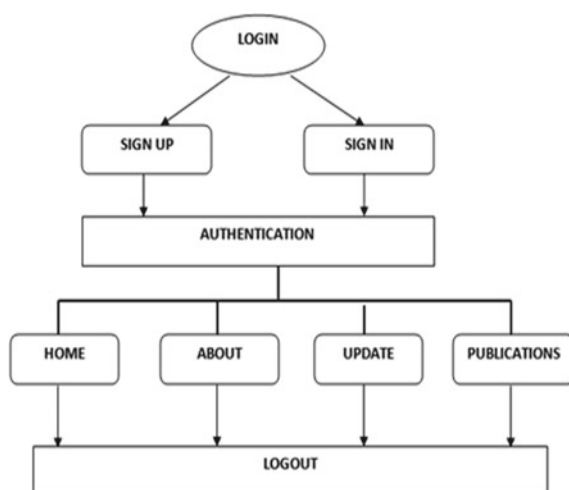
This is simple process wherein the user credentials (username and password) on the website's login form entered by the user are sent to the server for verification and authentication.

On successful completion of the authentication process, the user is granted with the permission to access their accounts. If the user is not registered, they can register themselves by filling up the registration form provided in the login page. The visitor's information is stored in the database.

#### 3.2 Home Page

The home page is displayed after the authentication process. This page gives a brief summary of the logged in user with track history. The home page displays the user

**Fig. 1** Workflow of the application



data such as the user’s ID, first name, last name and Email-id. It also displays all the publications submitted by the current user. The user is provided with facilities to edit or delete his existing publications. The CHANGE PASSWORD button is also available in home page.

3.3 About Page

About page displays the information about the administrators of this application and some details about the organization.

3.4 Publications Page

Displays the count of entire publications tabulated along with department wise segregation and allows the user in generating a customized report on clicking the view publications button at the left corner. The publications page has a VIEW PUBLICATION button at the bottom of the page. Clicking on this button displays the report generation page (Fig. 2).

The users can generate advanced reports with fast intelligent filters and multiple parameters provided. The user is provided with five filtering options to extract the particular data from the database. A process in which a subset data is obtained from a particular database to refer and analyse is Data filtration [7]. The filtering process is made so simple such that even non-tech users can retrieve data with a few clicks.

The results based on the user’s parameters, will be displayed in table format with records placed consecutively row-wise. If there are no matching records ‘No records found’ message will be displayed. The data flow of the application is shown in Fig. 3.

- The processor goes through each of the command to check which is similar to the user given filtering options.
- When the match is found, the program executes the corresponding dynamic SQL queries on the SQL database using the raw function of python.

Fig. 2 Report customizing options (view publications page)

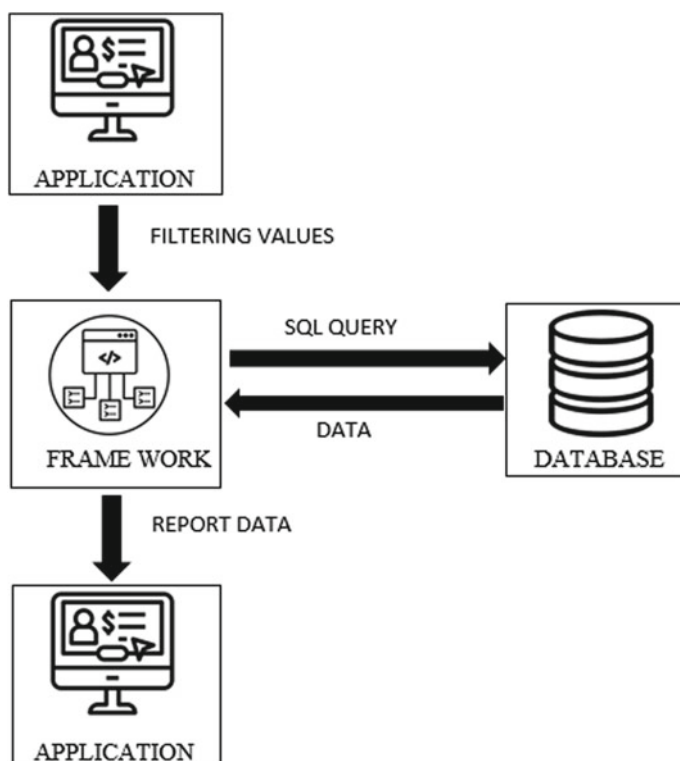
DEPARTMENT :

SOURCE :

ACADEMIC YEAR:

CALENDAR YEAR:

AUTHORS :



**Fig. 3** Data flow in the application

- The filtered records are returned as a query set to be displayed in the application.
- The SQL queries are dynamic since the variables are assigned with values based on the user's criteria. The dynamic SQL queries makes it reusable and also makes the code efficient and simple.
- After all the records are returned back to the processor, the retrieved data from the database is presented in a predefined format to the user on the interface (Fig. 4).

### 3.5 Update

The update page is used for data entry. This page consists of the update form with all fields of the table in the database to be entered by the user. After filling in all the required details and on clicking the UPLOAD button, the form validation process takes place. During the validation process, if the data entered by the user is found to be existing in the database, the user is alerted with a pop-up message "This publication already exists". Thus data duplication is avoided. On successful validation, the inputs of the user is uploaded into the database. The data collection process is made simpler

S.No	Publish Year	Document Title	Authors	ISSN	Journal Title	Month	Volume	Issue	Source	Department
1	2014	A five-factor software architecture analysis based on far for ATM banking system	Rathinababu T.K.S. Sankaranam N.	19928645	Journal of Theoretical and Applied Information Technology	May	59	3	SCOPUS	CSE
2	2014	A framework for safe composable testing model for multiple applications testing environment	Smitha P.S. Sankaranam N.	19928645	Journal of Theoretical and Applied Information Technology	May	63	2	SCOPUS	CSE
3	2014	Discriminative human action recognition using HOI descriptor and key poses	Alaka K. Chitrakala S.	-	2014 International Conference on Science Engineering and Management Research, ICSEMR 2014	February	-	-	SCOPUS	CSE

Fig. 4 Report data displayed on the interface

and there will be no need for an administrator to regularly collect and update database manually. Thus the data collection process is automated.

4 Database Architecture

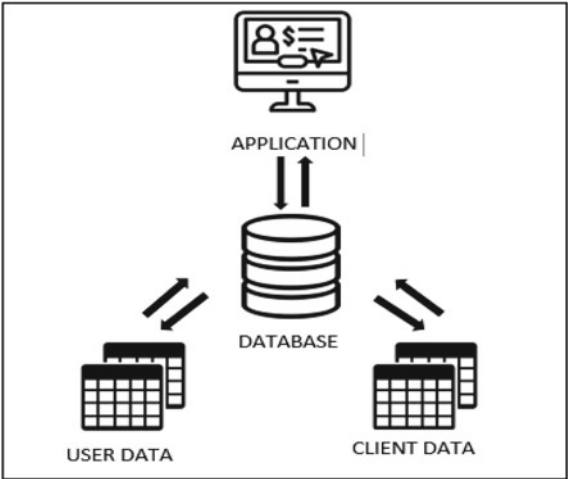
The database consists of two tables. The first table consists of the user profile data. The second table consists of the actual data which is updated by the users and used by the program in generating records. The type and quantity of data is not important for this application because with minor configurations, it can be made to work with any kind of database. The first table contains general information of the user and his permissions and access rights. This table is used for the authentication process as it has the passwords stored in hashed format. The second table contains all the necessary documents and previously entered data that can be used for report generation (Fig. 5).

5 Report Generation

A customized report is the finished product of this application. For a report to be generated, filtering options of the user is taken as an input. Based on the conditions given, the application works with the program to find the right match of criteria. When the conditions are met, the previously written SQL query for the particular selection gets executed in the database. The returned records are stored in a variable and passes to the application. All the records are completely imported onto the application and displayed to the user.

Once all the returned records are displays the user is provided with two formats to download the customized report. It can either be downloaded in pdf format which will be useful in saving space, easier sharing and produce hard copies. The other option is to download the report as an excel file for editing (Figs. 6 and 7).

Fig. 5 Database architecture



The screenshot shows a PDF report with a table containing the following data:

S.No	Publish Year	Document Title	Authors	ISSN	Journal Title	Month	Volume	Issue	Source	Dept
1	2014	A five-factor software architecture analysis based on far for ATM banking system	Rathishbabu T.K.S., Sankaran N.	19828645	Journal of Theoretical and Applied Information Technology	May	50	3	SCOPUS	CSE
2	2014	Discriminative human action recognition using HOG descriptor and key poses	Akila K., Chitrakala S.	-	2014 International Conference on Science Engineering and Management Research, ICSEMR 2014	February	-	-	SCOPUS	CSE
3	2014	A framework for safe composable testing model for multiple applications testing environment	Smitha P.S., Sankaran N.	19828645	Journal of Theoretical and Applied Information Technology	May	63	2	SCOPUS	CSE

Fig. 6 Report in PDF format

The screenshot shows an Excel report with a table containing the following data:

S.No	YOP	Document Title	Authors	ISSN	Journal Title	Volume	Issue	Month	Source	Department
1	2014	A five-factor software architecture analysis based on far for ATM banking system	Rathishbabu T.K.S., Sankaran N.	19828645	Journal of Theoretical and Applied Information Technology	50	3	May	SCOPUS	CSE
2	2014	Discriminative human action recognition using HOG descriptor and key poses	Akila K., Chitrakala S.	-	2014 International Conference on Science Engineering and Management Research, ICSEMR 2014	-	-	February	SCOPUS	CSE
3	2014	A framework for safe composable testing model for multiple applications testing environment	Smitha P.S., Sankaran N.	19828645	Journal of Theoretical and Applied Information Technology	63	2	May	SCOPUS	CSE

Fig. 7 Report in excel format

## 6 Advantages

- The main advantage of this application is its interoperability. It can be customized to work with any kind of database can be used in various government, private enterprises and institutions.
- The size and type of the database will not cause a major effect on this application. For large databases, with necessary hardware the proposed system will function without any errors [8].
- The application is user friendly and does not require any technical knowledge in handling the website.
- The application is intelligent enough in identifying the updation of an already existing data, to avoid duplicacy and hence saves space.

## 7 Conclusion

The main goal of the presented program is to simplify the three major processes of data handling such as data collection, data filtration and report generation. The application is useful in managing vast data and generating various customized reports from them. Provides advanced report retrieval system, can be downloaded as a pdf or export them as spreadsheets through an automated process with very less human effort.

It allows the user to manipulate the database by adding in new data through the provided webpage, by doing so they can not only add in new data but also generate a customized report immediately, right after adding.

This system can be implemented in most of the places wherein large data needs to be gathered, stored, analysed and customized reports generation are to be automated for speed and accuracy. For example, it can be used in government sectors, private enterprises, in educational institutions to maintain student records, in organizing patient's medical records, in warehouses and accounting firms for various tasks etc. With minor changes made, the application works on any kind of database, thus making it interoperable and widely compatible. The users of this application will interact only with its interface, a web-based GUI (Graphical User Interface). It is made of several web pages allowing the user to login to the application, access all the available webpages (HOME, ABOUT, PUBLICATIONS, UPDATE) and can download a customized final report.

## References

1. Benfield JA, Szlemko WJ (2006) Internet-based data collection: promises and realities. *J Res Pract* 2(2)



2. Sethuraman R, Kerin RA, Cron WL (2005) A field study comparing online and offline data collection methods for identifying product attribute preferences using conjoint analysis. *J Bus Res* 58:602–610
3. Malhotra R, Bansal A, Jajoria S (2016) An automated tool for generating change report from open-source software. In: International conference on advances in computing, communications and informatics (ICACCI), pp 1576–1582. <https://doi.org/10.1109/ICACCI.2016.7732273>.
4. Cooper CJ, Cooper SP, del Junco DJ et al (2006) Web-based data collection: detailed methods of a questionnaire and data gathering tool. *Epidemiol Perspect Innov* 3:1
5. Grunske L, Whalen M (2018) Guest editorial: advanced topics in automated software engineering. *Autom Softw Eng* 25:743–744. <https://doi.org/10.1007/s10515-018-0245-6>
6. Hyun C, Park H (2019) Image recommendation for automatic report generation using semantic similarity. In: International conference on artificial intelligence in information and communication (ICAIIIC), pp 259–262. <https://doi.org/10.1109/ICAIIIC.2019.8669018>
7. Mommessin C, Dreher M, Raffin B, Peterka T (2017) Automatic data filtering for in situ workflows. In: IEEE international conference on cluster computing (CLUSTER), pp 370–378. <https://doi.org/10.1109/CLUSTER.2017.35>
8. Ray HS, Mukherjee S, Mukherjee N (2020) Performance enhancement in big data handling. In: International conference on contemporary computing and applications (IC3A), pp 17–22. <https://doi.org/10.1109/IC3A48958.2020.233261>

# On Trade Cloud Ecosystem Structure for Shared Learning



Vaibhav E. Narawade and Vishal S. Badgujar

**Abstract** Cloud computing (CC) is a latest type of network automation stack; its advancement velocity is growing in a very increasing fashion. In this work, we proposed an implementation of cloud ecosystem for shared learning to explore and use infrastructure, platform and software as a service (IaaS), (PaaS) and (SaaS), which makes a prerequisite planning and design of the general environment of learning cloud for the needs of education domain. With this system, workload of a staff can be reduced with respect to laboratory setup for various laboratory subjects in each semester of an academic year. It allows to access cloud over Web for shared learning, so as to call the required application programs with respect to various subject groups together as and when required.

**Keywords** Cloud · IaaS · SaaS · PaaS · XEN · Ulteo · Hypervisor · Shared learning

## 1 Introduction

Cloud computing is nothing but the storage on net servers instead of on the system [4]. It is just like having an addition of storage in terms of hard drive that one can use anywhere, anytime  $24 \times 7$  connected to the network which gives different computations through different vendors of cloud. So, whenever we are utilizing this cloud type for the computation service meant one can accept input and process it and provides return back to customer is called as the CC.

Over network, computation parameters that are reached as a service or application is nothing but the cloud computing [3] which is provided by most of the service vendors such as Amazon Web Services, popularly known as AWS, Microsoft Azure platform and GCP.

---

V. E. Narawade (✉)

Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India

V. S. Badgujar

Department of Information Technology, A. P. Shah Institute of Technology, Thane, Mumbai, India

The cloud computing application models are basically contributed into the three ways which are (a) Infrastructure (IaaS), (b) Platforms (PaaS) and (c) Software (SaaS). It has large domains of applications; today's plenty new methodologies coming in public domain, so by using this, one can identify certain automations and check it as and when we needed.

## 2 Service Models

Service models are classified broadly as: (a) Infrastructure, (b) Platform and (c) Software.

IaaS: It gives us with infrastructure components such as virtualization, disk, networking and computation. This type of service model we are not having any basic control over on this cloud setup, but one can use it for deploy and also run software services utilizing it [2]. Its purpose is file backup that is restoration and processing for limit of time and product design. E.g., for such service you can consider (AWS)-Amazon Web services.

PaaS: It gives us with virtual machine environment advancement. A client can design and launch different softwares on the cloud environment. One not having control on this infrastructure but having on its deployed services or software. E.g., for such service, you can consider (GCP)-Google App Engine.

SaaS: It [6] gives access to service vendors' services or applications that execute on this infrastructure or platform. This will allow multiuser or clients to use an application on the network in order to use services readymade. Users can use this feature with own devices like BYOD concept remotely using a Web browser and a GUI [4]. The users cannot be able to control the infrastructure features of cloud, such as computing power, net, servers, OS and disk, except for few configuration settings will be required. E.g., Gmail, Google Docs and Drop Box.

## 3 Collaborative Learning

It is a method of teaching and learning which includes different students group working all together to resolve a common problem facing, complete a task or create a final solution or product. When we talked about collaborative learning in terms of cloud framework for education field, we must understand the scenario before cloud and after our proposed solution for smooth implementation of cloud services.

Before suggesting the cloud for collaborative learning, we will understand one of the problems in education field: in each semester for every class, there is an need to setup lab manually for desired softwares which will be required for teaching in that specific semester's subject, but here, the problem is the subject incharge or lab incharge needs to setup labs with required software and pre-requisite for every new batches and for every new lab as labs and subject are not the same for each faculty.

In order to do that by using SaaS on demand service of cloud computing, we are able to use it to overcome this issue or problem and give collaborative hassle-free platform for faculty members and students to do lab exercises without wasting time for installation and configuration; they can directly use on demand applications or softwares by just clicking on icons of software which is already installed on their login.

One of the main advantages of this system is every user or students what we are targeting got separate logins or workspace with personal authentication as well to maintain personal data, lab outputs, logs and so on. Let us understand why such system is required; e.g., students who are doing lab exercises complain that in one lab session, whatever they performed and saved output or documents that are not found in next lab session the reason is that other subjects' lab practical also scheduled in that particular lab and might be other students deleted that files, so in order to avoid this, we will give all students a different separate platform to perform practicals which contain all subjects' required software applications installed and configured so that students get an readymade environment to directly work upon in terms of on demand software as a service of cloud which is nothing but here is an collaborative learning.

## 4 Xen Server

XenServer which will run or work on your machine, so there is nothing between your machine and XenServer service. Naming XenServer will work as the OS. It will also communicate to machine hardware part along with NIC-network interface card and machine storage. It requires a CPU with 64-bit architecture with virtualization that is Intel VT or AMD-V enabled on machine hardware. Now, this specification will be more like an acceptable specifications. It does not mean that XenServer is not going to install on your system if CPU does not have Intel Virtualization Technology ON, it will be worked. But the functionality of virtualization will be very limited in such case. XenServer is a para virtualized type-1 hypervisor. It uses the para virtualization type of a technique of server virtualization one.

It needs the system processor that you have on physical server to be capable of doing such a type of virtualization in order to provide its capabilities. We need at least 2 GB of memory of machine to install XenServer, but the suggested is at minimum must have 8 GB of RAM, especially when intend to host multiple VM's on the top of it. XenServer will nearly 16 GB of storage. But, in this case, we will require more for our virtual machine's setup. It can be work with a NIC Controller of 100 MB. However, what is acceptable is a gigabit controller. Setting up XenServer is easiest and fast too.

1. You can use a CD by Burning the ISO to a CD and use it into drive for installation.
2. Also needs to enable virtualization technology in your systems BIOS on which you need to install.
3. Then you can boot to the XEN bootable CD.

4. Whatever questions on display answer it and reboot.
5. Make initial configurations, like allocating an IP to machine, a hostname, and password for security.
6. After that you can connect to your XenServer using a client XenCenter (For Windows) or OpenXen manager (For Linux) with respect to your OS which you have.

## 5 Proposed Methodology

To solve the problem, we use the Ulteo OVD software service delivery base for education domain cloud [1]. That is nothing but the Ulteo 2.0 was introduced as a kind of open virtual desktop. Ulteo platform is based on Deb architecture and Ubuntu which enable users from multiple devices through the Web browser from Unix and Win software's too. With respect to Ulteo application, the customers require to active Jvm-based Web browser to run whatever applications you wanted to use. The UNIX software's bunch includes many applications. Also, one can install Linux applications via the management utility. Importantly, in this, one can get the support of RPC of software's in windows part. It will give possible constraints for the most clients which are familiar with the windows services. The Ulteo basically having three main parts: Ulteo Apps server (worked on both UNIX and Windows), Ulteo SM server and Admin part.

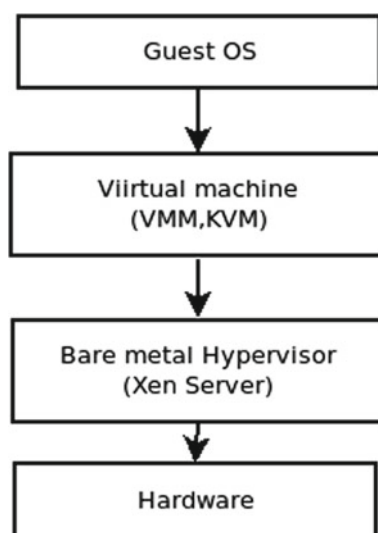
In Fig. 2, we will use server virtualization [5] to deploy our infrastructure as a service that is XenServer and host our software as a service Ulteo, Bare metal hypervisor which is a type-1 hypervisor, for example, XenServer in our case what we are using here. Bare metal hypervisor is directly installed on hardware; there is no operating system on hardware as shown in Fig. 1. We will install XenServer on hardware by using suitable client application, and we will get the access of this XenServer in order to install guest operating system on XenServer.

In our case, we will install Ulteo as an on demand application delivery software as a service (SaaS) on XenServer as a IaaS as shown in our proposed methodology which gives you direct access to the required software applications. Ulteo is only the virtualization technique or solution which will support both Linux and Windows Dual applications; Ulteo open virtual desktop is also called as OVD.

**Case Study:** In this proposed methodology, we implemented one case study on college laboratories which includes creation of different classrooms students' group together and enrolled or registered them into the Ulteo OVD admin panel by accessing user tab and provide them an authentication separately for security. For example, SE, TE and BE three classes with 60 students each registered into their respective groups. Then, we will find out the subjects from different semesters and sort out the application which is required for laboratory session subject wise.

After identifying the subjects and respective software applications, we will form a group of that applications subject wise and year wise mapped together and assign to particular group of classes. In order to use the same software version through the

**Fig. 1** Xen architecture bare metal (type-1) hypervisor



class at the same time and same infrastructure without installation and configuration of any kind of applications. So, when user or students' needs to perform practical or lab exercise, he or she just needs to access client portal of OVD and provide authentication which is setup by faculty from admin panel earlier, and after successful login, students can get all software applications within their login itself.

## 6 Xen Center and Open Xen Manager

In order to deploy our Ulteo OVD on infrastructure as a service XenServer, we will require a client program depending on operating systems for communication because Xen is directly installed on the hardware. For the Windows operating system, we require a Xen center as a client program and for Linux OpenXen manager. Installation of Xen center is easy just like any other software you need to download and double click and the next button needs to press until it is installed. Now, we will see how to install OpenXen manager on Linux operating system.

Installation of OpenXenManager in Linux.

Step 1. Firstly, ensure Linux system packages are updated by performing apt-get commands on terminal as:

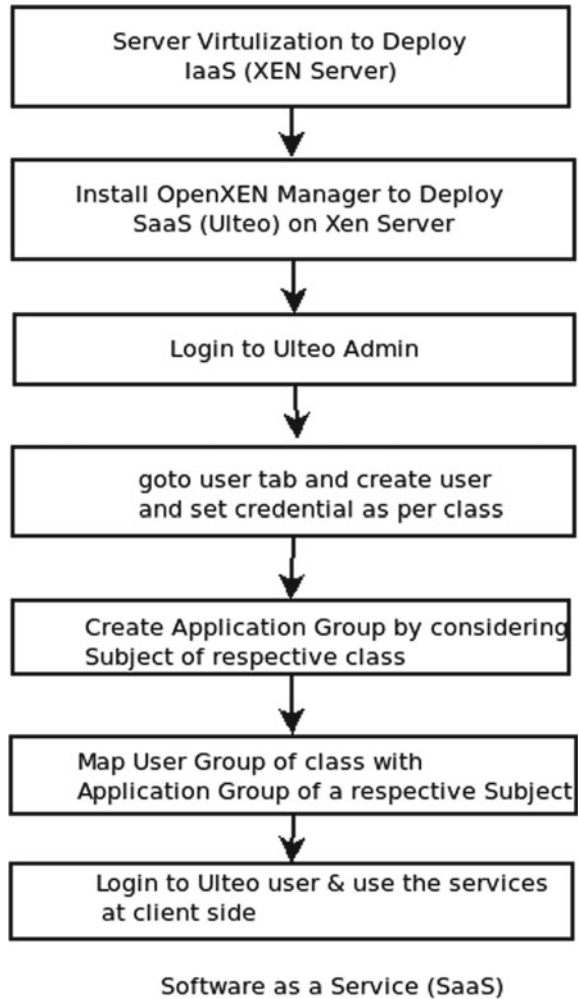
```
vaibhav@vaibhav:~$ sudo apt-get upgrade
```

```
vaibhav@vaibhav:~$ sudo apt-get update
```

Step 2. Installing package of python and all dependencies for installation of OpenXenManager through terminal as:

Perform below commands to install the packages on your system:

**Fig. 2** Proposed methodology



```
vaibhav@vaibhav:~$ sudo apt-get install python-gtk2 glade python-gtk-vnc
python-glade2 python-configobj
```

Step 3. Install Openxenmanager.

```
vaibhav@vaibhav:~$ git clone https://github.com/openxenmanager/openxenmanager.git
```

Switch from your existing folder to cloned repo directory through terminal:

```
vaibhav@vaibhav:~$ cd openxenmanager
```

after this install the package by performing below command as given:

```
vaibhav@vaibhav:~$ python setup.py install
```

Step 4. Access OpenXenManager on your Linux platform.

Done with the installation; if one want then you utilize software on system, and for that, just run the below command as: vaibhav@vaibhav:~\$openxenmanager.

## 7 Deployment of a Software as a Service (SAAS)

For installation of client program, you need to connect it with XenServer and deploy our Ulteo [7, 9] on it by creating virtual machine of Ulteo on XenServer using OpenXen manager.

Step 1: Install Ulteo through DVD or Open Ulteo OVF file in Xen Center by selecting import on Xen Server [8].

Step 2: At Management machine Open following URLs.

<https://Ulteo-Server-ipaddress/ovd> for Client access

<https://Ulteo-Server-ipaddress/admin> for Admin access

Step 3: Login on Admin portal specify Username and password as admin.

- Under server tab Register server, click on manage to add IP address of Ulteo Server.
- Go to user tab to add multiple users as we discussed in case study multiple users from different class SE, TE and BE.
- Go to Application Tab to Create Application Group for specific subject of respective semesters software applications which shows in Figs. 3 and 4.

Then go to User Tab Create to User Group as a SE, TE and BE.

Map User group with Application group and use the services at client side.

Step 4: At client side open <https://Ulteo-Server-ipaddress/ovd> for Client access, Specify Username and Password for personal workspace as shown in following Fig. 5.

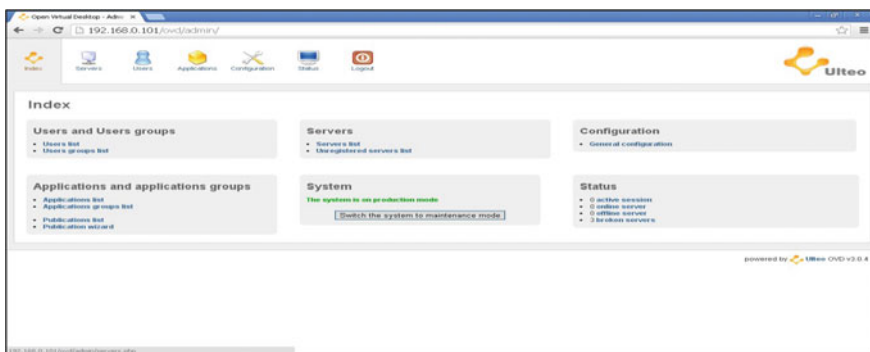


Fig. 3 Ulteo admin portal (server side)



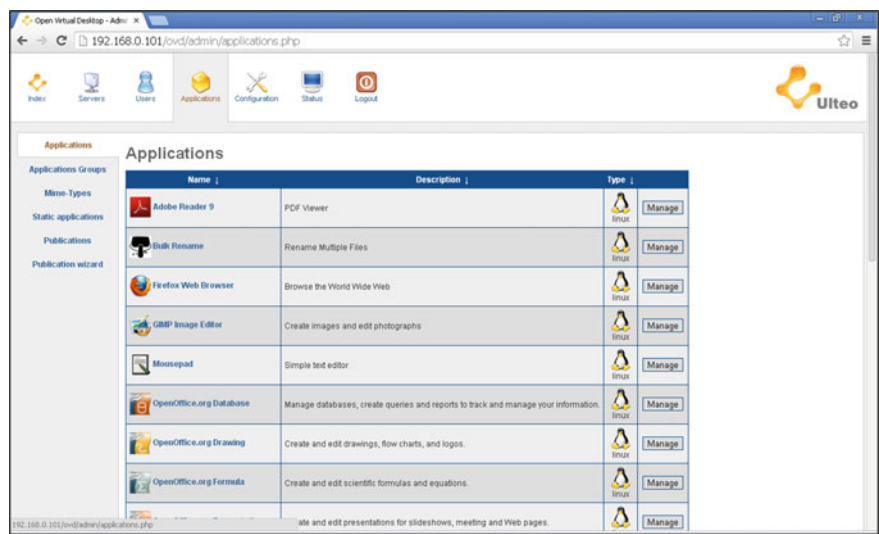


Fig. 4 Ulteo application group configuration

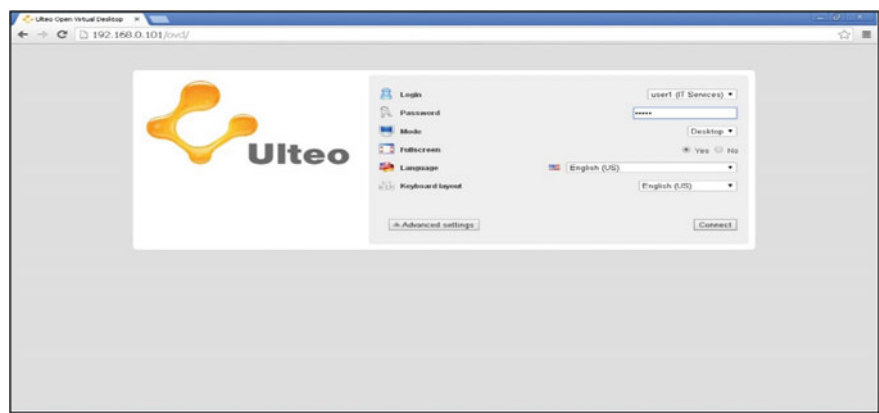


Fig. 5 Ulteo user portal (client side)

After successful login, the client portal which having all required software for respective semesters for laboratory exercises which we called as on-demand application delivery, which is a software as a service deployment-ready platform for direct use the software applications (Fig. 6).



**Fig. 6** On-demand application delivery SaaS (Ulteo virtual desktop)

## 8 Conclusion

In this work, we built a virtual on-demand desktop system using Ulteo OVD successfully which will rely on process of CC, of application deployment in Edu domain infrastructure for collaborative learning-based research. By using this platform, we can work spreading out more in the laboratories instead of installing softwares every time. We also discuss about the cloud, its services and types; it is nothing but a resource sharing or communicating way over the Internet. With separate user space and data security in SaaS, we were also hoping that this new technique or advancement can be taken over to the education domain for better education service.

## References

1. Chen WD, Wang Q, Ye XD, Xu YF (2012) A study of cloud education environment design and model construction. In: 2nd International conference on Consumer Electronics, Communications and Networks (CECNet). <https://doi.org/10.1109/cecn.2012.6202308>
2. Prajapati Amit G, Sharma Shankarlal J, Badgujar Vishal S (2018) All about cloud: a systematic survey. International conference on smart city and emerging technology (ICSCET). Mumbai, India, pp 1–6
3. Hayes B (2008) Cloud computing. *Commune ACM* 51(7):9–11
4. Mell P, Grance T (2011) The NIST definition of cloud computing
5. Pandya V, Navageni V (2016) Cloud computing strategy: cloud server and architecture. *Int Educ Res J* 2(4)
6. Zhu Z, Yang G, Zhou F (2011) The application of SaaS based cloud computing in the university research and teaching platform. In: International conference on intelligence science and information engineering. <https://doi.org/10.1109/isie.2011.19>
7. Velte AT, Velte TJ, Elsenpeter R (2009) Cloud computing: a practical approach. 68–73
8. Ulteo Open Virtual Desktop <https://distrowatch.com/table.php?distribution=ulteo>
9. Linux and Windows Virtual Desktop <http://www.ulteo.com>

# A Comparative Study of Local and Global Feature Detector and Descriptors for Image and Video Retrieval



Ekta Sarda, Jayshree Jain, and Vaibhav Narawade

**Abstract** In this paper, we are studying about two leading approaches. Step one, we will define the local features detection and description method for classifying pattern recognition, image tracking, clustering object recognition, or identification images based on local image patches or key points in the images of an object. Secondly, we can define global feature detection and description methods for an image that are used in image retrieval and object detection for a whole image. Local and global feature methods are comprehensive, and their output is compared using a dataset. The images were selected according to the hypothesis that they could be better described using global features. Selecting the algorithms, that all depends on the feature's detection and description for the data. This led to a detailed review of local and global video or image detectors and new descriptors. We presented a summary of existing performance evaluation and databases of benchmarks. Lastly, in terms of future directions, we have completed the investigation. This research study may assist and provide a framework for image recognition/processing and machine learning.

**Keywords** Local feature detector and descriptor • Global feature detector and descriptor • Matching pattern • Benchmark datasets • Video copy detection

## 1 Introduction

A feature is an element of information that is relevant to solving the calculation task related to a particular application. Characteristics may be specific structures

---

E. Sarda (✉)

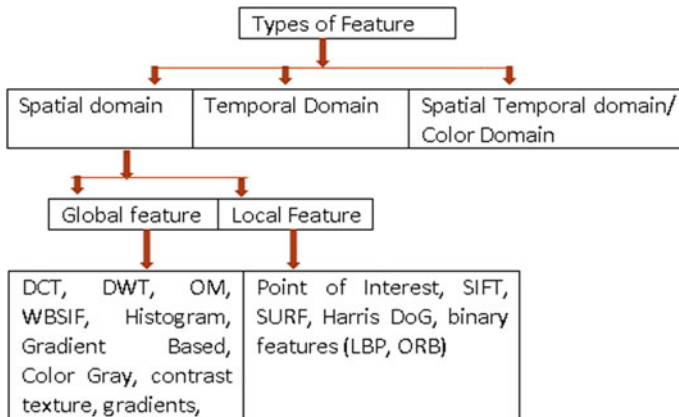
Faculty of Computer Engineering, Pacific Academy of Higher Education and Research University, Udaipur, India

J. Jain

Department of Computer Engineering, Pacific Academic Higher Education and Research University, Udaipur, India

V. Narawade

Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India



**Fig. 1** Types of features

within the image such as dots, edges, or objects. Features can be classified into three dimensions, and further, spatial domain can be classified into two types as shown in Fig. 1.

- I. **Spatial dimension:** Each image in a pixel space lies in a different area, and the dimension extracts the color distribution or layout of objects within the frames [1], such as ordinal signature and gradient orientation centroids.
- II. **Temporal dimension:** The sequence of images represents the key frames based on their dimensions, movement detection, etc. [1], are popular examples of the exploitation of time information.
- III. **Color dimension:** It depends on the color or gray level properties of the frames, such as histogram, tint, saturation [2].

### 1.1 Feature Extraction

In pattern recognition and image processing, a feature retrieval is a special form of dimensional reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g., the same measurement in both feet and meters), then the input data will be transformed into a reduced representation set of features. Zhang et al. [3] Transforming the input data into the set of features is called feature extraction. Feature extraction involves simplifying the number of resources required to describe a large set of data accurately.

Features extraction techniques are used to reduce various datasets such as PCA, ICA, LDA, and LLE [1, 4].

### 1.1.1 PCA

Principle component analysis (PCA) is used for unsupervised learning; it means that it is used for high variance and low obstacle for dimensionality to minimize the error of reconstruction.

### 1.1.2 ICA

Independent component analysis (ICA) is to eliminate all the unnecessary noise. ICA is a linear method of reducing dimensionality which integrates as input information of different data mixture and aims at correctly identifying them. Whether their linear or not linear dependency is equal to zero, two input functions may be considered independent. Its application is used for medical as well as image processing.

### 1.1.3 LDA

Linear discriminant analysis (LDA) is to increase the distance between the mean of each class and reduce the spread within the class. It is based on supervised learning for reduction techniques and classification of machine learning. Its disadvantage is that it is not working in Gaussian distribution.

### 1.1.4 LLE

Locally linear embedding (LLE) is based on nonlinear features, and its reduction techniques are based on manifold learning to design an object with high dimensional space instead of unnecessary space. It is different from PCA and LDA.

Feature extraction methods are of three types defined as follows: spatial domain, temporal domain, spatial and temporal domain; other methods are local and global feature.

1. **Spatial domain:**—Wary and Neelima [1] It is based on feature vector extracted from local and global of a video such as local feature used for lossy compression, resizing, frame rate change, scaling, and rotation, and in global feature, overall frame dropping and logo insertion function is not used in spatial domain since it takes up a large memory space for the exploitation of the video/ image database.
2. **Temporal domain:**—it is used to differentiate time and frame. The functions and the hash value are extracted between two consecutive video images to acquire the image level representation internally. The temporal domain only uses the content relationship in the time activity phase to its global descriptor cannot be used for transformation to change the subsection frame of the video, i.e., cropping and frame insertion of a large area.

3. **Spatial and Temporal Domain:**—Wary and Neelima [1] it gives the better performance results for all shortcoming and their challenges. It uses the 3D-DCT and 3D-RBT video sequence. Video frame contains time information for moving video and certain distortion based on contrast enhancement and frame dropping such as picture in picture and frame insertion.

## 1.2 Feature Transformation and Selection

Azhar and Thomas [2] describe the data transformation to increase algorithm accuracy. Changing existing features to create new ones from the old ones. The key factor analysis that uses an orthogonal transformation is a more widely used technique for dimension reduction, such that there are linear unrelated variables dependent on the original collection of variables. Features selecting that have the greatest impact on the target variable as well as existing features can be done with various techniques such as linear regression, decision trees, calculation of importance weights (e.g., Fisher score, ReliefF), and removing unnecessary features.

## 2 Local Features

Local feature exploits just parts of the whole image, pattern of images, keypoint frame that differs from its immediate environment. In contrast, local features identify points of interest from frames (such as edges, corners, or blobs), and then describe the local region surrounding those points. Joshi and Patel [4] In other words, local features seek abrupt changes in pixel intensity values and their relationship to neighboring pixels. Intensity, color, and texture are commonly considered to represent the output of a frame and to select keyframes within a sequence, eliminating redundancies while optimizing the coverage of particular key points in a keypoint set.

The local feature is used to change image matching, recognition of objects, 3D reconstruction, and location for pictures, picture rotation, zoom, light, and viewpoint. Their performance is robust and relies heavily on local stability characteristics. Its different example implements to calculate local feature descriptors such as SIFT, floating points such as SIFT, SURF ORB FREAK, and LBP.

### 2.1 Scale Invariant Feature Transform (SIFT)

Keypoint differences in the detection of local maximum values and those points of interest are defined according to gradient directives. Based on the good stability and discrimination, it is possible; Lowe [5] uses a SIFT descriptor. SIFT performs well

between categories of local characteristics and is robust to scale, rotate, rush, and affinity.

## **2.2 *Speeded-Up Robust Features (SURF)***

It is based on discrete wavelet responses summarized by the point of interest that give the Hessian determinants a maximum value [6]. SURF has great strength to deal with geometric shifts such as image scaling, translation, and rotation. For representing local max points of interest that has enhanced real-time efficiency.

## **2.3 *Hessian-Laplace Feature***

The Hessian affine detector and Laplacian of Gaussian are combined in this characteristic. He uses the Laplacian Gaussian to find invariant points on several scales. In each scale, the maximum value is reached both for the trace and for the Hessian determinant [7].

## **2.4 *Oriented FAST and Rotated BRIEF***

Joshi and Patel [4] describe the ORB and SIFT perform the feature detection task (and better than SURF) with a faster magnitude of nearly two orders. ORB builds on a popular BRIEF descriptor and FAST key point detector. The good performance and low costs of both these technologies are attractive. The major contributions of ORB are: (i) add a rapid and robust FAST guidance part. (ii) Effective BRIEF-driven function calculation. (iii) Analysis of BRIEF-oriented variance and correlation. (iv) A method of learning to decorate BRIEF is rotationally invariant, resulting in better results in the nearest applications.

The challenge of local characteristics is to ensure a robust description of key points. Different robustness and complexity of local features vary. Most descriptors are intended to be invariant to scale, rotate, and thinly transform, but not invariant against severe camcording, reencoding, and flip changes, to mention a few. Although floating vector descriptors are robust against many graphical variants, binary descriptors require significantly less space at a very small descriptive ability loss.

The advantage of local descriptors is that they only focus on some points of interest rather than the whole image, so even if the background which occupies a big space has been changed, as long as the signatures only exploit the information of key points, we can still detect the copies of the local features. Local features have a good performance on geometric–theoretical framework transformations and edition [2, 3, 5, 6]. However, they tend to have large computing costs for extraction and matching.

### 3 Global Feature

Global feature is called as global descriptor, global descriptor is generally used in image retrieval, object detection and classification. It includes contour representation shape, descriptor, and texture features. Video transformations are used for feature invariant such as such as compression, change of contrast/brightness/gamma, noise addition, geometrical transformations (such as rotation, affine transformations, scaling, etc.); post production or edition (PiP, insertion of patterns, captions, pillar boxes, etc.). FLIP is horizontal mirroring, CAM is camcording, and DES refers to desynchronization attacks (temporal, spatial). Global features can significantly improve the survival rate, and the visual accuracy of the key points can be improved.

All the entire video sequence the global features are determined, copies of all the sequences may be found, but some specific operations such as logo insertion and cropping. Global descriptors use techniques based on the temporal domain, spatial domain, and spatial–temporal domain. Shinde and Chiddarwar [8] define global feature extraction techniques categories are as follows: PCA form is used for the whole face, LDA method is used for Fisherface, different transformations are Gabor filter, Waveletsub bands, and optical matrix factorization, ICA method, DFT (discrete Fourier transform), DCT (discrete cosine transform), DWT (discrete wavelet transform), spectrogram, chronogram, and so on.

#### 3.1 Histograms and Edge Histogram

Zhao [4] define that a histogram is a gray or color histogram representing the picture. The gray intensity histogram reflects the gray level distribution in the picture, discarding the spatial places. The descriptor of the edge histogram reflects the spatial edge distribution in the picture [6].

#### 3.2 Ordinal Measurement

Shinde and Chiddarwar [8] capture the relative order of intensities in the picture. The technique of extraction transforms an image into a gray scale and splits it into nanometer areas. The average intensity for each area is calculated. The ranks express the relative ordering between areas. Ravi and Gowda [7] define the function is used to find a color loss correspondence in final images, but it was not resilient to changes such as rotation. This function provides an ordered sequence of video frames focused on their average ability. For the generation of video segment's signature, Zhang et al. [3] explained that in the video, frame is divided into block numbers than the average gray value is determined for each block.



### **3.3 DCT**

In removing redundancy within adjacent pixels, it is essential to use picture conversion. In the reconstruction of an image, quantizers are also allowed to remove coefficients at small amplitudes without temperature differences [8]. DCT has a low-frequency energy transfer, which decreases the total amount of data needed to explain a video or a frame of an image. Azhar and Thomas [2] Extracting the DCT-Sign-based function by applying two dimensional discrete cosine transform to each block predetermined the keyframe for AC-coefficients is used as a source and search keyframe-signature.

### **3.4 DWT**

Wavelet transformation is used to obtain video frames for functional descriptors. The video is split into a number of parts with a fixed length of frames. Multilevel wavelet decomposition the function vector is obtained by using these frames. Thampi [4] A representative vector for each segment is obtained by taking the mean of the consecutive vector for each segment. A video cluster is created and a cluster center, calculating the mean of the representative vector, is calculated to facilitate the similarity search. In two phases, search occurs: (i) the distance between the query representative vector  $q_i$  is calculated and the clusters center  $Cr$  is calculated and the winning video cluster is achieved. (ii) Calculate the distance between  $q_i$  and the vector of the segments that make up the selected cluster Cluster  $Cr$ . Calculation wavelets allow the computation of compact function or data representations. The choice of discrete wavelet transformation is clarified by attaining varying degrees of information or precision, computational attractiveness, etc.

### **3.5 Color Based Feature**

Shinde and Chiddarwar [8] define features-based signature is easy, but it is sensitive to changing colors. As color transformations are the most common assault on copying videos and color signatures, the light or gray image is used for most machine copying of all the HOG binaries in each of the sub-regions of the video images. However, luminance-based approaches do not work well for transformations like trimming, pillar box, and text insert.

## 4 Motion Features

The use of motion features that are moving events in a video series, as they remain unchanged in extreme deformations, will efficiently solve this problem. Thaipanich [5] the content-based video copy recognition, the motion feature did not find the best solution that are listed here—(i) If the movement is captured frequently at frame rate, there is almost no movement activity, so no valuable information can be recorded. (ii) Motion vectors retrieved from normal imaging rates may look miscalculated in any manner as the neighboring pixel sizes are identical one another in successive video frame. (iii) There is not much motion for a static video material such as a news channel interview program, which means that the vector value is virtually minimal low.

## 5 Temporal Features

Temporary domain are changes in the artifacts in relation for the frequency domain instead of considering the spatial nature of each video frame. Shooting length series is distinctly robust, as there are no successive anchor frames of identical timescales for any different video sequences.

## 6 Both Global and Local Feature

Local and global feature defines to retain spatial robustness and discriminatory properties of the functional descriptors for a video framework similarity preserving for compact hash code generation. This method was used to extract from each keyframe local and global feature, Wary and Neelima [1] Using the SIFT and Pyramid histograms of the directional gradients (PHOG) to combine the two features to build a small compact hash code with good precision.

## 7 Descriptor Performance Evaluations

Table 1 describes that recall and precision measure the performance evaluation of the features.

**Detector Performance Evaluation:** the essential properties of local and global characteristic for all evaluation are ruggedness, repeatability, accuracy, simplicity, and quantity. The criteria are widely used in numerous applications as vital criteria for assessing the local and global function detector. Joshi and Patel [4, 7] Apart from

**Table 1** Performance evaluation of the features

Precision	TP/ TP+FP	Precision represents a relevant fraction of cases (accuracy). It provides an accurate ratio of one between the number of false matches and the total number of matches	TP stands for true positive, and FN stands for false negative
Recall	$R = TP/TP+FN$	Recall is the number of regions that are correctly matched for the number of regions on the same two-frame scene	TP stands for true positive, and FN stands for false negative
F1 score	$F\text{-Score} = 2PR/P + R$	F1 score is calculated using precision and recall. It provides the system with overall performance	P stands for precision, and R stands for recall

the different geometric transformations, it evaluates the ability of the detector to extract the same paying turns.

**Repeatability rate** =  $D/\min(L1, L2) * 100$  Where  $D$  is number of corresponsive,  $L1$  is number of pictures in first frame and  $L2$  is number of pictures in second frame.

## 8 Benchmark Datasets Table

In Table 2, we analyze the performance evaluation metrics, applicable algorithm, datasets, and the suitable application that are used in combination of local and global descriptor and detector as well.

## 9 Conclusion

This paper aims to provide the current image or video detection method with a detailed description. Many of the methods available to detect images, and video copies that is the combination of spatial, temporal, and all temporal domains based on the characteristics derived, and many more techniques were used. Many researchers have developed methodologies that utilize both spatial and temporal information in a video to resolve these issues. In the last few decades, the identification of copied or pirated copies of original video content was compounded by the huge advent of multimedia technology. The use of the methods which have to discriminate and robust against different material conservation and geometric properties. Depending on data type, the dataset size and the application the algorithm is highly selected.

**Table 2** Prior work on performance assessments of detectors and descriptors of local and global characteristics

Descriptor and detector	Performance evaluation metric	Algorithm (feasible solution)	Application	Benchmark datasets
Local descriptor [1, 8, 5, 6, 3]	Euclidean distance matching	DAISY	Image matching	TRECVID 2010
	F-score	LIOP, RFD	Image matching	MUSCLE-VCD2007
	Repeatability information content	Harris	Geometric transforms	TRECVID2003
	Feature density, entropy density	MSER	Image coding	MUSCLE VCD
Global detector [8, 4, 7]	High repeatability in segment-level matching	DCT	Invariant keyframe and scaling	TRECVID2003
	Euclidean distance+clustering-based search	DWT, Spatial correlation descriptor	Image classification	Fashion-MNIST
Global detector+Local detector [1, 3, 4, 7]	SIFT+ordinal measure	Transformation adaptive matching	Geometric Transformation	KTH, UCF-sports

The video copy detection method has been rendered most complicated by attacks like loss compression, resizing, rotation, scaling, etc. Numerous researchers are now working on problems to improve its efficiency and performance.

## References

1. Wary S, Neelima A (2018) A review2 on robust video copy detection. *Int J Multimedia Inf Retrieval*
2. Azhar MA, Thomas PA (2019) Comparative review of feature selection and classification modeling. In: 2019 International conference on advances in computing, communication and control (ICAC3), 2019
3. Zhang H, Zhao Z, Cai A, Xie X (2010) A novel framework for content-based video copy detection. In *IEEE proceeding of IC-NIDC*, 2010
4. Joshi K, Patel MI (2020) Recent advances in local feature detector and descriptor: a literature survey. *Int J Multimedia Inf Retrieval* 1–17
5. Lowe DG (2004) Distinctive image features from scale invariant keypoints. *Int J Comput Vision* 91–110
6. Bay H, Tuytelaars T, Van Gool L (2006) SURF: Speeded Up Robust Feature”, in *Proc. of European Conf. on Computer Vision*, Springer LNCS, volume 3951, part 1, pp. 404–417, 2006

7. Ravi C, Gowda RM (2020) Development of image stitching using feature detection and feature matching techniques. In: 2020 IEEE international conference for innovation in technology (INOCON)
8. Shinde SR, Chiddarwar GG (2015) Recent advances in content based video copy detection. In: 2015 International conference on pervasive computing (ICPC)

# Extracting Attributes of Students Mental Health, Behaviour, Attendance and Performance in Academics During COVID-19 Pandemic using PCA Technique



Vinayak Hegde, M. Shilpa, and M. S. Pallavi

**Abstract** Corona virus disease (COVID-19) outbreak affected schools, universities, and colleges across the world to an unscheduled shutdown of the campuses so that students can maintain social distancing measures. However, moving effortlessly from a traditional educational environment to distance and virtual learning could not happen overnight. The primary motivation of the study is to analyse and understand the effectiveness of online education and its complications encountered by the students to use ICT devices and requirement of Internet connectivity issues, economic crisis, emotional and distress management, distractions, anxiety and confusion about the studies and health issues among students. In our study, an overall of 224 students delivered complete information concerning the survey. Data were collected through structured questionnaires administered in four different sections to acquire the perspective about the online mode of learning from both undergraduate and postgraduate students of different colleges with the help of Google Form through WhatsApp. Each section determines the factors that are affecting students in academics, behaviour and mental issues, and performance gives a statistical description of online learning during the pandemic. With these data collected, we are going to analyse and establish a statistical analysis and find the correlation between 30 factors using principal component analysis (PCA) to estimate the learning approach, learning status and numerous issues related to online learning due to lockdown. As a result, we obtained 16 positive PCA mean values as truly supported and 14 negative values as not supported hypothesis carried to predict whether students are suitable to attend online class or not in future analysis.

**Keywords** Student academic performance • Mental health • Behaviour • Attendance • Educational data analytics • Principal component analysis (PCA) • Hypothesis • Dimensionality reduction • COVID-19 pandemic • Online learning

---

V. Hegde (✉) • M. Shilpa • M. S. Pallavi

Department of Computer Science, Amrita School of Arts and Sciences, Amrita Vishwa Vidyapeetham, Mysuru, India

## 1 Introduction

The World Health Organization (WHO) confirmed COVID-19 as a global public health emergency of international concern on 30 January 2020 along with a first-phase nationwide lockdown on 22 March 2020 in India. The COVID-19 cases were increasing rapidly all over the world. In response to the pandemic, Indian experts shut all the educational organizations across the nation on 13 March 2020 so that students could maintain social distancing measures. According to the guidelines given by the Government of India (Pragyata Guidelines), the Higher Education Commission India (HECI) and University Grants Commission (UGC) gave mandates to higher education organizations to begin planning and arranging for distance learning (DL) approaches, reschedule the undone exams and help their students consistently online until the COVID-19 emergency stays unaltered [1]. In accordance, traversing easily from traditional education to distant online learning environments could not occur instantaneously. This rapid change of environment is interconnected to numerous challenges and obstacles.

The primary aim of the study is to analyse and to spot the effectiveness of online education among students in their academics based on attendance, behaviour, mental health, and performance and various difficulties faced by the students to use ICT devices and requirement of Internet connectivity, economic crisis, distractions, anxiety and confusion about the studies and health issues among students are the main objectives.

The sample of study includes the student's understanding level of concepts during the online classes, around 14(6.3%) of students were clearly understanding the concepts, 90(40.2%) of the students understands moderately and 19(8.5%) of students understanding level is very poor. Statistics say that 96(42.9%) of students say that they were very much affected by the connectivity issue, 36(16.1%) of students say that moderately affected and only 16(7.1%) students say that they were not much affected.

With the collected data from the survey, we are analysing and establishing a statistical analysis along with the hypothesis testing using the dimensionality reduction technique known as principal component analysis (PCA) to estimate the student's learning status, issues in the learning approach, student perspective about online learning and other difficulties related to online study due to the pandemic lockdown.

Finally, considering hypothesis test results, the positive results of the hypothesis were considered to be supported and the negative results of the hypothesis were considered to be not supported were carried for the further process of analysis to predict students comfort level in attending distance online learning classes.

## 2 Literature Review

COVID-19 pandemic has stuck the conventional and habitual learning method of all the academic organizations and pushed the management of schools, universities and colleges to opt for virtual learning as a diverse way to restart education. This abrupt change results in facing different kinds of challenges by undergraduate and postgraduate students and examining their attitudes towards online education and its effectiveness. In this research study, demographic data were acquired through a structured online survey using the Likert scale and were analysed and stated in percentages by common frequency of students' responses [2]. To examine the implications of coronavirus that lead to school closures on the education, using STATA/Regression analysis interpretation, factor analysis, distribution of respondents with the attributes [3] and to know how effectively shifting of the educational structure of institutions can transform the existing resources of properly structured schooling into virtual education with the help of virtual classes [4].

With the fast advancement of ICT devices and the proportionate experience of students to it inspects students' views about online learning which will help in acquiring learning skills and it enhance their performance with a better understanding [5], which gives a comprehension of how online learning tends to unsuccessful virtual learning systems and anxiety of year loss in academics influences learner's mental wellness [6].

Virtual learning was discovered to be beneficial as it provides more ease for the students, but also for some students, virtual learning could be more complicated than regular habitual learning due to the technical limitations, late feedback and the incapability to handle the information effectively [7].

The statistics were analysed to compare the results with the current to previous results. Also, if the grade information is available, then it will be compared with previous results in both traditional and online classes to obtain the relative results [8]. The extent of COVID-19 has radically interrupted each part of human life including the educational sector. Because of the COVID-19 pandemic, numerous new methods of learning, new patterns and new perspectives were emerged [9].

The unique COVID-19 pandemic corroborated by unique EMA data makes the results somewhat difficult to compare to previous works, for the mental health measures identified a decrease in loneliness [10]. The e-learning stages carried out likewise represent a test to most of the learners because of restricted usage of the Internet and the absence of specialized information on these advanced gadgets [11]. The COVID-19 worldwide outbreak provided an opportunity to discover deep information about virtual learning [12].

Results show that there is an increment in pressure and anxiety because of the COVID-19 flare-up. Different stressors were perceived that added to the expanded degrees of stress, nervousness and burdensome musings among learners [13], and also findings depict that there is a greater occurrence of mental health disputes, which are emphatically aligned with social media contact during the COVID-19 outbreak



[14]. The exhibition of the learners can be estimated on various measures utilizing Z-scores and scores are put in their dispersions [15].

### **3 Methodology**

#### ***3.1 Data Collection and Procedure***

A structured survey was conducted online using the Google Form link from 10 November to 25 December 2020 to gather information about the student's attendance, behaviour, mental health and performance in academics. An overall of 224 students conveyed total and genuine data concerning the survey. Descriptive statistics and simple percentage distribution were accessed to evaluate the learning grade, learning approach, student's attitudes towards online education, mental health issues and other difficulties linked with the study during the lockdown. Every analysis was performed using the R language [16].

#### ***3.2 Research Model and Hypotheses***

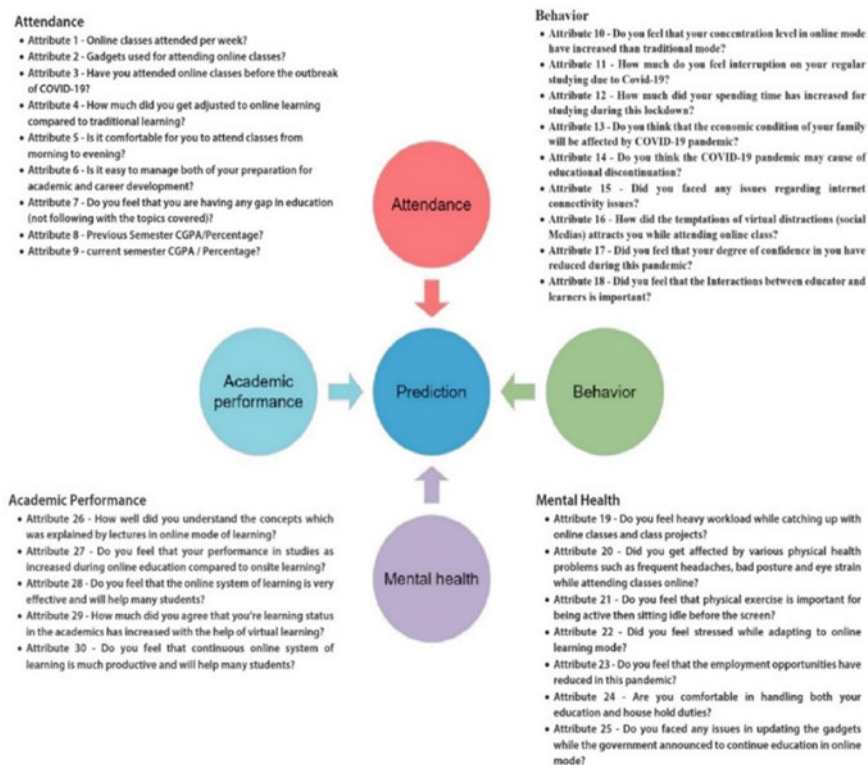
This study is an approach to analyse the student's attendance, mental health, behaviour and performance in academics. As shown in the Figs. 1 and 2 represent the research model includes 30 attributes like student's concentration level, mode of the online class, economic condition and other attributes considered for analysing student's attendance, behaviour, mental health and performance in academics, and these are the four-parameters considered for the analysing student's overall performance.

##### **3.2.1 Attendance**

Attendance is the model in student's academics which helps them know the state of going regularly to class. Students' class attendance and their commitment play a vital role in modern higher education and predicting student's attitudes towards learning [17]. Students who attend classes regularly have been shown to achieve at higher levels than students who do not have regular attendance [18].

##### **3.2.2 Behaviour**

Behaviour is one of the important features in knowing the level of learning that the student has. Effective learning wants behaviour that lets you fully participate in the lesson, and motivate other students to do work as well.



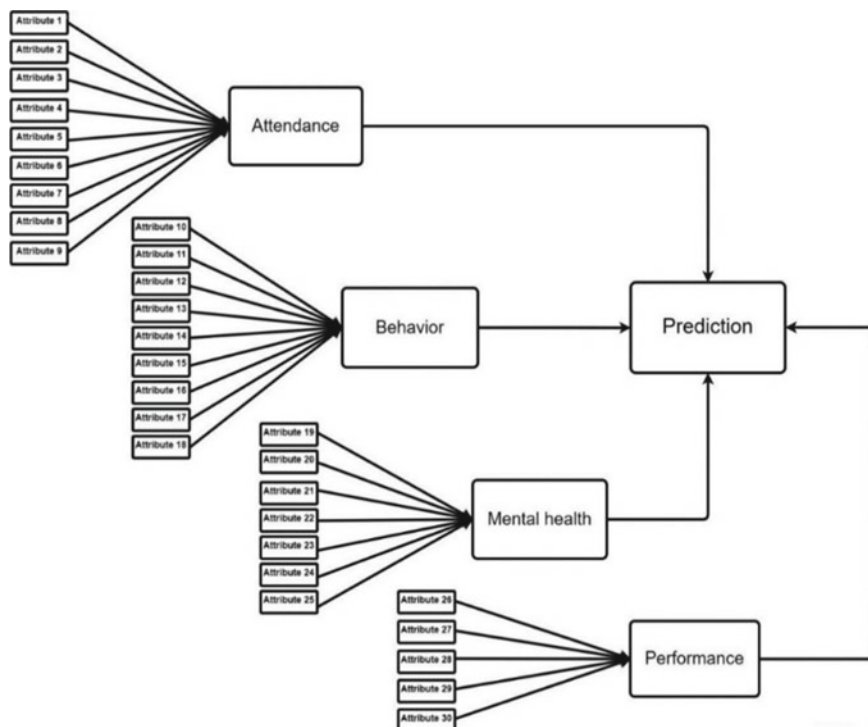
**Fig. 1** A research model for analysing students' performance through attendance, behaviour, mental health and performance in academics

### 3.2.3 Mental Health

Mental health is one of the vital features in knowing their mental stability to learn new things and update themselves. There are multiple factors involved. It can influence a student's energy level, focus, reliability, positive thinking, mental capacity and obstructing performance. Over the past few years, students' mental health has been deteriorating because of academic stress and anxiety [19].

### 3.2.4 Performance in Academics

The academic performance of a student is the amount of student achievement across various academic subjects [20]. Students who improve themselves in school are better able to make the progress into adulthood and achieve occupational and monetary progress.



**Fig. 2** A conceptual model for analysing students' performance through attendance, behaviour, mental health and performance in academics

### 3.3 PCA

Principal component analysis (PCA) is a dimensionality reduction technique of a large dataset, while conserving as much information (variable) as possible. In principal component analysis (PCA), we either be able to choose the matrix of covariance or the matrix of correlations for components. In our study, we have opted for the covariance matrix given in supplementary material. After standardizing the data, PCA tries to get the features with maximum variance from the data using a covariance matrix to get the eigenvalues and eigenvectors. If the element of the covariance matrix is positive, then the attribute's value will increase or decrease together (correlated variables). If the element of the covariance matrix is negative, then one attribute's value increases when the other decreases (inversely correlated variables).

4 Results

After standardization of data, principal components are to be extracted using Eigenvalues and Eigenvectors with maximum variance with the help of covariance matrix. The scree plot and the variance bar plot graphs are shown in Fig. 3a, b, respectively. The principal component PC1 vs PC2 is plotted as shown in Fig.3c. In our research results, it shows that the first two principal components have a high range of variances in the data set with a variance of 18% (PC1) and 12% (PC2), respectively as shown in Fig. 4a, b. We evaluated the mean value of the corresponding PCA to determine supported and not supported hypothesis results. The positive mean value

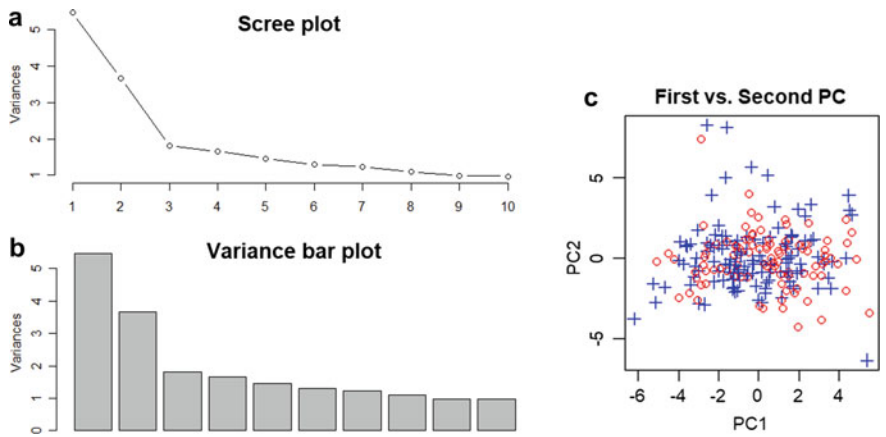


Fig. 3 a Scree plot, b variance bar plot, c plotting principal components

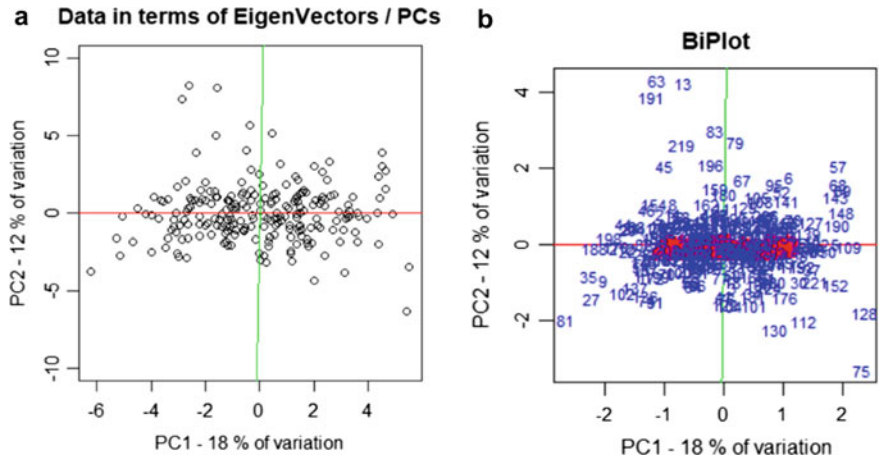


Fig. 4 a Data in terms of EigenVectors/PC's b BiPlot of PC's

of PCA shows that the hypothesis determined is supported, and the negative mean value of PCA shows that the hypothesis determined is not supported. Thus, supported hypothesis variables results can be taken for further process of analysis.

## 5 Discussion

This research paper mainly concentrates on choosing the appropriate attributes from an online survey, for analysing student's attendance, mental health, behaviour and performance in academics during the COVID-19 pandemic, conducted among students from various colleges and universities. The survey consisted of thirty important parameters from attribute 1 to attribute 30 which exhibits the student's perspective on virtual learning, as shown in the research model (Figs. 1 and 2). Through the survey, we could able to gather 224 data from the students, which includes different sections of structured questionnaires to estimate the learning status, opinions on educational decisions, mode of learning and various problems associated with the study due to the lockdown. Data are displayed with the scale and loadings for attendance, behaviour, mental health and behaviour in supplementary materials. Then, using those data Tables, we will apply principal components to calculate the Eigenvectors and Eigenvalues using the matrix of covariance of the data and also data in terms of eigenvectors or principal components (PC's) plots are shown in Fig. 4a, b. Using the obtained PCA values, let us calculate the mean value of PCA to obtain the result of the hypothesis. The positive mean value of PCA shows that the hypothesis determined is supported, and the negative mean value of PCA shows that the hypothesis determined is not supported for the analysis. Thus, supported hypothesis variables results are taken for the further step to analyse student's attendance, mental health, behaviour and performance in academics during the COVID-19 pandemic. As a result, we obtained PCA mean values as 16 positive and 14 negative values are represented in supplementary materials. This paperwork is to distinguish which parameters are truly supported and which are not supported for further process of analysis. In which, PC1, PC10, PC11, PC13, PC15 and so on have the highest PCA mean values as shown in Table 1, which represent these variables are highly supported for the prediction.

## 6 Future Work

The attributes related to attendance, behaviour, mental health and academic performances are hypothetically tested with the help of principal component analysis (PCA), which is known to be a dimensionality reduction technique. Finally, considering hypothesis test results, the positive results of the hypothesis were considered to be supported, and the negative results of the hypothesis were considered to be not supported attributes. With the obtained result, the supported (positive mean)

**Table 1** Demonstrates the supported attributes of PCA in order

PC's	PCA mean value	Hypothesis name
PC1	0.061815	Online classes attended per week
PC10	0.022956	Do you feel that your concentration level in online mode has increased than the traditional model
PC11	0.00086	How much do you feel interruption on your regular studying due to COVID-19
PC13	0.015133	Do you consider that your family's economic circumstance will be affected by the COVID-19 pandemic
PC15	0.022592	Did you face any issues regarding Internet connectivity issues
PC17	0.001211	Did you feel that your degree of confidence in you have reduced during this pandemic
PC19	0.031096	Do you feel a heavy workload while attending online classes and class projects
PC20	0.001599	Did you get affected by various physical health problems such as frequent headaches, bad posture and eye strain while attending classes online
PC21	0.003636	Do you feel that physical exercise is important for being active than sitting idle before the screen
PC22	0.00593	Did you feel stressed while adapting to online learning mode
PC26	0.011158	How well did you understand the concepts which were explained by lectures in the online mode of learning
PC27	0.001795	Do you feel that your performance in studies has increased during online education compared to onsite learning
PC30	0.010386	Do you feel that a continuous online system of learning is much productive and will help many students
PC5	0.009529	Is it comfortable for you to attend classes from morning to evening
PC6	0.010528	Is it easy to manage both of your preparation for academic and career development
PC7	0.028549	Do you feel that you are having any gap in education (not following with the topics covered)

PCA attributes can be taken for the further process of analysis to predict whether the student is suitable in attending distance online learning classes or not with their upcoming academic results.

## 7 Conclusion

The COVID-19 epidemic has caused worldwide learning systems, leading to the overall terminations of schools, universities and colleges, which lead to the start of educating students online with the help of the virtual platform for all the educational

sectors. This study helps in understanding and demonstrating the solutions to the numerous effects of COVID-19 on students' education.

This paperwork is to distinguish which parameters are truly supported and which are not supported for further process of analysis. In doing so, we can diminish the components of a huge informational collection, by altering a large set of attributes into a modest one that contains paramount information in the huge set.

### Supplementary materials

<https://drive.google.com/drive/folders/11FTmmqQ6cTNNZQSHq66yX71MdmJf-EhO?usp=sharing>

## References

1. Kapasia N, Paul P, Roy A, Saha J, Zaveri A, Mallick, R, Barman B, Das P, Chouhan P (2020) Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India. *Children Youth Serv Rev* 116(June). 105194 <https://doi.org/10.1016/j.childyouth.2020.105194>
2. Adnan M (2020) Online learning amid the COVID-19 pandemic: Students perspectives. *J Pedagog Res* 1(2):45–51. <https://doi.org/10.33902/jpsp.2020261309>
3. Onyema EM, Sen S, Alsayed AO (2020) Impact of coronavirus pandemic on education. *J Educ Practice* (June). <https://doi.org/10.7176/jep/11-13-12>
4. Mishra L, Gupta T, Shree A (2020) Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *Int J Educ Res Open* 1(September), 100012 <https://doi.org/10.1016/j.ijedro.2020.100012>
5. Kalpana R, Mahajan, MV (2020) A study of students' perception about e-learning. *Indian J Clin Anat Physiol* 5(4):501–507. <https://doi.org/10.18231/2394-2126.2018.0116>
6. Hasan N, Bao Y (2020) Impact of “e-Learning crack-up” perception on psychological distress among college students during COVID-19 pandemic: a mediating role of “fear of academic year loss.” *Children Youth Serv Rev* 118(August):105355 <https://doi.org/10.1016/j.childyouth.2020.105355>
7. Muthuprasad T, Aiswarya S, Aditya KS, Jha GK (2021) Students' perception and preference for online education in India during COVID-19 pandemic. *Soc Sci Humanit Open* 3(1):100101 <https://doi.org/10.1016/j.ssaho.2020.100101>
8. Basilaia G, Kvavadze D (2020) Transition to online education in schools during a SARS-CoV-2 Coronavirus (COVID-19) pandemic in georgia. *Pedag Res* 5(4). <https://doi.org/10.29333/pr/7937>
9. Madeshia PK, Verma S (2020) Review on higher education in India. *J Crit Rev* 7(10):1161–1164. <https://doi.org/10.31838/jcr.07.10.228>
10. Fried E (2020) Mental health and social contact during the COVID-19 pandemic: an ecological momentary assessment study. <https://doi.org/10.31234/osf.io/36xkp>
11. Owusu-Fordjour C, Koomson CK, Hanson D (2020) The impact of COVID-19 on learning-the perspective of the Ghanaian student. *Eur J Educ Stud* 7(3):88–101. <https://doi.org/10.5281/zenodo.3753586>
12. Son C, Hegde S, Smith A, Wang X, Sasangohar F (2020) Effects of COVID-19 on college students mental health in the US: an interview-survey study (Preprint). *J Med Internet Res* 22:1–14. <https://doi.org/10.2196/21279>
13. Wargadinata W, Maimunah I, Dewi E, Rofiq Z (2020) Student's Responses on learning in the early COVID-19 Pandemic. *Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah*, 5(1):141–153. <https://doi.org/10.24042/tadris.v5i1.6153>

14. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, Dai J (2020) Mental health problems and social media exposure during COVID-19 outbreak. *PLoS ONE* 15(4):1–10. <https://doi.org/10.1371/journal.pone.0231924>
15. Hegde V, Pallavi MS (2016) Descriptive analytical approach to analyze the student performance by comparative study using Z score factor through R language. In: 2015 IEEE international conference on computational intelligence and computing research, ICCIC 2015, pp 0–3. <https://doi.org/10.1109/ICCIC.2015.7435813>
16. Hegde V (2017) Dimensionality reduction technique for developing undergraduate student dropout model using principal component analysis through R package. In: 2016 IEEE international conference on computational intelligence and computing research, ICCIC 2016, pp 1–6. <https://doi.org/10.1109/ICCIC.2016.7919670>
17. Hegde V et al (2020). Predicting students performance through behavior and computational thinking in programming. In: Tuba M et al (eds), *ICT Systems and sustainability, advances in intelligent systems and computing*, p 1270, [https://doi.org/10.1007/978-981-15-8289-9\\_40](https://doi.org/10.1007/978-981-15-8289-9_40)
18. Fadelelmoula T (2018) The impact of class attendance on student performance. *Int Res J Med Med Sci* (June):47–49. <https://doi.org/10.30918/irjmms.62.18.021>
19. Aguilera-Hermida AP (2020) College students' use and acceptance of emergency online learning due to COVID-19. *Int J Educ Res Open*, 1(September):100011 <https://doi.org/10.1016/j.ijedro.2020.100011>
20. Devasia T, Vinushree TP, Hegde V (2016) Prediction of students performance using educational data mining. In: *Proceedings of 2016 international conference on data mining and advanced computing, SAPIENCE 2016*, pp 91–95. <https://doi.org/10.1109/SAPIENCE.2016.7684167>



# Emotion and Gender Classification Using Convolution Neural Networks



Chiranjeevi Pandi , K. Adi Narayana Reddy, Ramesh Alladi,  
V. Chandra Sekhar Reddy, and P. Sumithabhashini

**Abstract** Gender classification and detection of emotions play an important role in the field of security. Gender classification classifies the person's gender into male or female, whereas emotion detection helps in identifying the feeling of the person. Several machine learning algorithms were proposed for the classification of gender and detection of emotions individually. In this paper, we proposed a mechanism to detect emotion and gender classification simultaneously from face images. The deep learning algorithms were efficient over machine learning algorithms on large amount of data. Convolution neural networks is one of the hierarchical deep learning architectures used for classification and detection of images. So, in this paper, we used a convolution neural networks for classification of gender and detection of emotion. The experimental results demonstrate that the proposed mechanism outperformed over the traditional machine learning algorithms.

**Keywords** Machine learning · Deep learning · Classification · Detection · Convolution neural networks · Gender · Emotion

## 1 Introduction

Development of big cities and towns and rapid urbanization has been contributing to a larger extent to the crime rates. The phenomenal rise in offences and crimes has become a greater concern for the security and safety of the population. As the technology has been invading into most of the domains, we make an attempt to adopt the mechanism of gender and emotion classification using convolutional neural networks through deep learning technique into crime which enables identification

---

C. Pandi (✉) · R. Alladi · V. Chandra Sekhar Reddy  
ACE Engineering College, Hyderabad, India

K. Adi Narayana Reddy  
BVRIT Hyderabad College of Engineering for Women, Hyderabad, India

P. Sumithabhashini  
SAMSKRUTI College of Engineering and Technology, Hyderabad, India

and classification of faces of previous criminals and helps in crime detection and prevention in a minimal amount of time without delay through this work.

In recent years, we had come through many CNN models. Every new emerging CNN model come up with better improved features of the previous models. CNNs are mainly used for image classification. The birth of CNN took place in 1998; till now, we have approximately 10 mainly used CNN models. Some of the models according to their published years are LeNet-5, AlexNet, VGG, Network in Network, Inception-v1, ResNets, Inception-v3, Xception, Inception-v4, Inception-ResNets, and ResNeXts. Each model has their importance according to their use and need in the different applications.

The proposed module consists of two modules; they are emotion and gender classification. The input is given as image which goes through these modules. Firstly, it classifies face from the given input, and then, the classified face goes through the other two models. Emotion classification means, identifying the facial expression of a person. Humans can make many expressions, out of which the mainly seen expressions are happy, sad, angry, disgust, neutral, fear, and surprise. The emotion module recognizes the emotion of a person by classifying the facial expressions. Gender classification means, identifying the gender of a person based on the parameters taken from face.

This module only deals with the two genders, male and female. The output is given in the form of text on the top of the face present in the image (e.g., Happy man or neutral woman).

The Deep Learning Technique: CNN model used in this architecture is Xception with depth-wise separable convolutions, and the type of CNN used is sequential. The models published before the Xception has many combinations of convolutions such as  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$ . These convolutions are replaced with depth-wise separable convolutions. The main feature of Xception is that it purely deals with the depth-wise separable convolutions. We choose Xception because its size is 88 MB, Top-1 Accuracy is 0.79, depth with 126, and best part is that it takes less number of parameters [23 M] when compared to other models. More quantity of parameters can decrease the robustness of the architecture.

In both the modules, we use Xception with depth-wise separable convolutions. The depth-wise separable convolutions consists of two, depth-wise convolution and pointwise convolution. Grad-CAM Technique is used in this module which highlights the important and relevant regions in the input image. This is used for transparent visualization which helps in easy dealing of the inputs. We are using this for better understanding of visual image classification and image captioning.

Inference methodology is used in this module to gain speedier and accurate results when new data is given. Inference cannot happen without training the data. Untrained neural network model is trained with the database which results in a trained neural network. This trained neural network is used for predictions of new data which takes less time to predict compared to the conventional training procedure.

## 2 Related Work

In research [1], they worked on static images to classify the emotion in faces. They used Fisher's linear discriminant (FLD), singular value decomposition (SVD), and radial basis functional network (RBF) for implementing, feature extraction, and network for classification, respectively. Only (happy, disgust, and angry) three expressions were able to classify with accurate values.

In research [2], the emotion recognition is done on both static images and video. It used facial animation parameter (FAP) and facial definition parameter (FDP) feature extracting and used only six emotion classes. It detects face and extract features defined on particular expression; by this, it obtained an accuracy of 60% on average. The more accurate results are obtained only on static images when compared to video input.

In research [3], the most popular classifiers BLSTM and HMM are used for expression classification. It says that bidirectional (BLSTM, BRNN) networks give more accurate than unidirectional networks. According to their work, neural networks give efficient results than the classification methods like HMM. At last, they got the accuracy of 61% by using various folds and clusters.

In research [4], they used large datasets which contains videos of people with mainly two emotions (amusement and sadness) which undergoes the linear and discrete classifications which includes SVM and LogitBoost classifiers to identify the emotion and gender of a person. This systems works only on few emotions, but there are many emotions a human can make. This is useful only to predict two emotions.

In research [5], the main agenda is to find the action units by using HMM and discriminant analysis. These action units can be used further to know the emotion of a human. It says that HMM with dense flow extraction gives more accurate results than the HMM with high gradient component detection. The more action units we consider, the more accuracy can be obtained.

In research [9], they have used VGGNet architecture of deep neural networks to classify the gender of a particular human by seeing the face. The steps in traditional VGGNet contains four levels, but in this work, they have reduced the step count from 4 to 1 and resigned it as their proposed architecture. It got accuracy of 0.85 for recognizing the gender.

In research [11], they have used only CNN with combination of dropout and augmentation techniques to classify the age and gender of the person. This method obtains an accuracy of 86% in detecting the gender. In research [13], they have used adaptive fuzzy network-based transfer learning for image classification, and they obtained of accuracy 88% for image classification.



**Fig. 1** IMDb dataset picture

### 3 Dataset

The datasets used in this work are IMDb and fer2013 for gender and emotion classification, respectively.

IMDb dataset stores the all the details of movies which includes each person details on who is part of the movie. It contains details of about 14,000 movies. As this dataset contains actors and actress with detailed perfections on their faces, it becomes easy to classify the gender of testing data by considering the IMDb dataset. The main reason to take this dataset for this work is that the data is preprocessed, kept ready for machine learning experiments, and updated every day.

Fer2013 dataset stores the images of people with different emotions; mainly, there are seven types of emotions present in it (angry, sad, disgust, happy, fear, surprise, and neutral). It is an open-source dataset which was used to a work, and later, it is published into the Kaggle. The dataset contains 35,887 grayscale images and 48X48 sized images. The priority of choosing this dataset is that it contains clear and perfect images with correct expressions in the faces (Figs. 1 and 2).

### 4 Proposed Model

In this work, we have analyzed many existing models and developed some new models by integrating the previous with new methods defined in it. By considering all these, we have developed two models. The two models help us to decrease the number of parameters in CNN last layer. There are many advantages of decreasing the parameter count; first, the slow performance can be eradicated by removing the small CNNs; the slow performance can be seen while using the hardware of low configuration. Second, the reduction of parameters increases the generalization process on Occam's razor framework. The first model idea is to remove the whole fully connected layers. The second model is the combination of deleting the fully connected layers and include residuals and depth-wise separable convolutions and both are trained under ADAM optimizer.



**Fig. 2** Fer2013 dataset picture

In first model, we took the help of global average pooling to delete all fully connected layers, and this worked out well such that the number of feature maps in last convolution layer is same as to the number of classes taken. Here, we also applied Softmax activation function for feature map which is reduced. This is standard fully convolutional neural network consisting global average pooling, batch normalization, ReLus and nine Layers. This contains approximately 6,00,000 parameters. The whole model is trained on both IMDB gender dataset and Fer-2013 emotion dataset; it achieved accuracy of 95% and 60%, respectively. This model is referred as “sequential fully-CNN.”

The second model is developed on the basis of Xception architecture; it uses both residual modules and depth-wise separable convolutions. Residual modules make modification between the two subsequent layer’s desired mappings, and then the difference of original feature map and desired feature map results in the learned features. The earlier learning problems  $F(x)$  are solved by changing the desired features  $H(x)$ .

$$H(x) = F(x) + x$$

In first model, the fully connected layers are removed by using global average pooling, but here, depth-wise separable convolutions are used. There are two layers in depth-wise separable convolutions: (i) depth-wise convolutions and (ii) pointwise convolutions. The main use of this is that it separates both spatial cross correlation and channel cross correlation. It also decreases the computation when compared with normal convolution layer. The difference between normal convolution layer and depth-wise convolutions are shown in below Figs. 3, 4 and 5.

Fig. 3 Standard CNN Filter

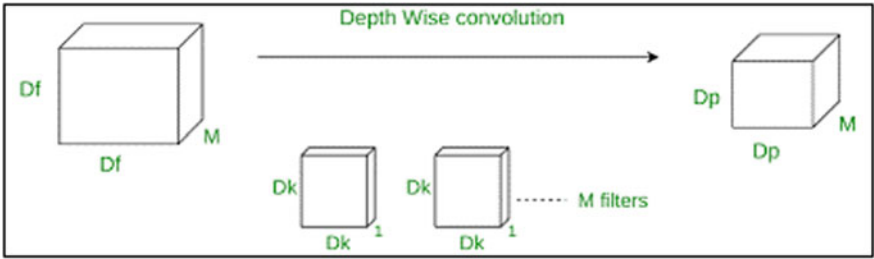
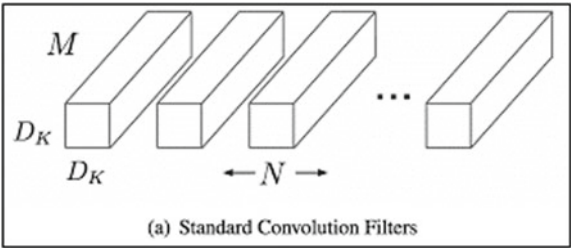


Fig. 4 Depth-wise convolution

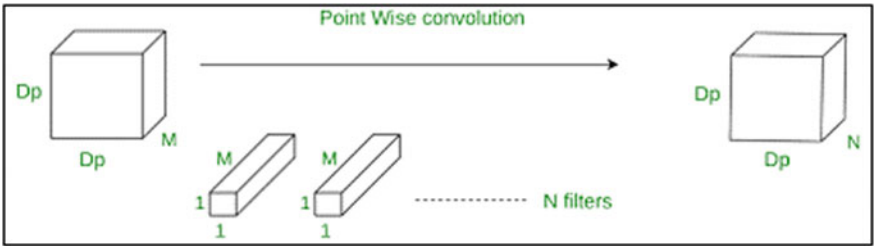


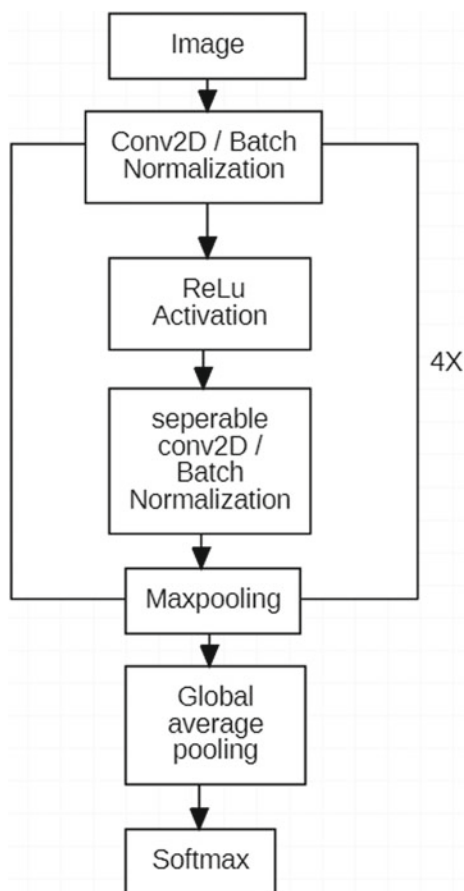
Fig. 5 Pointwise convolution

The final proposed architecture is fully CNN containing four residual depth-wise separable convolutions, ReLu activation layer. At last, Softmax activation and global average pooling is applied to get the results as prediction. This architecture has 60,000 parameters, whereas initial architecture contains ten times more parameters and original CNN contains 80 time mores parameters. Figure 6 shows details of proposed architecture which we named it as mini-Xception.

We used guided backpropagation technique in this to know which pixels or elements of an image are getting highlighted. This makes derivation of each element in input image with each element in the feature map. The result of derivation separates the negative gradients and the rest gradients are chosen because they only elevates the count of chosen elements in feature map.

The flowchart of the whole work goes like, the format of static input is.png file and dynamic input is taken through webcam which i.mp4 file. The input image is

**Fig. 6** Proposed architecture



converted into grayscale format then guided backpropagation is applied where we can decide the feature map for both gender and emotion classification.

The output of guided backpropagation is sent to both gender classification module and emotion classification module. The result can be seen in the output of respective models. The output is shown as two words, one is from gender class: {"Man," "Woman"} and the other word represents one of emotion from class: {"Happy," "Sad," "Angry," "Disgust," "Neutral," "Fear," "Surprise"}. The steps which takes place in modules are discussed in modules chapter.

## 5 Result Analysis

The whole pipeline of the work goes with detecting the face, classifying the emotion and the gender. This architecture achieved 96% accuracy in gender classification

**Table 1** Accuracy analysis of proposed and other existing architectures

Method used	Accuracy emotion classification	Accuracy of gender classification
Proposed system	76%	96%
HMM, BLSTM	61%	–
FAP, FDP	60% (only for 6 emotions)	–
FLD, SVD, RBF	45% (only for three emotions)	–
CNN with dropout and augmentation	–	86%
VGGNet	–	85%
SVM, LogitBoost	40% (only for small datasets)	50% (only for small datasets)

and 76% accuracy in emotion classification. By using the proposed architecture, we have reduced the computational cost and able to combine both emotion and gender modules such that the same input image is sent to both modules parallelly. The time taken by the system to detect is approximately 0.15 ms on i7-8550 CPU. This speed is more when compared to the original model developed by Tang. Many of the works contain either gender classification or emotion classification done on limited data which contains few number of figures, but this work contains both gender and emotion classification which runs at a time on huge datasets and gives accurate results.

The comparison of the proposed model and the existing models are presented in Table 1. The proposed model outperformed comparing with other existing models. The train and test accuracies are almost same, and there is no overfitting.

The results of this work is shown in Fig. 7, where the emotion classification output is shown taken from run-time input. The Fig. 8 shows the runtime output of both emotion and gender classification. The whole pipeline of the work goes with detecting the face, classifying the emotion and gender.

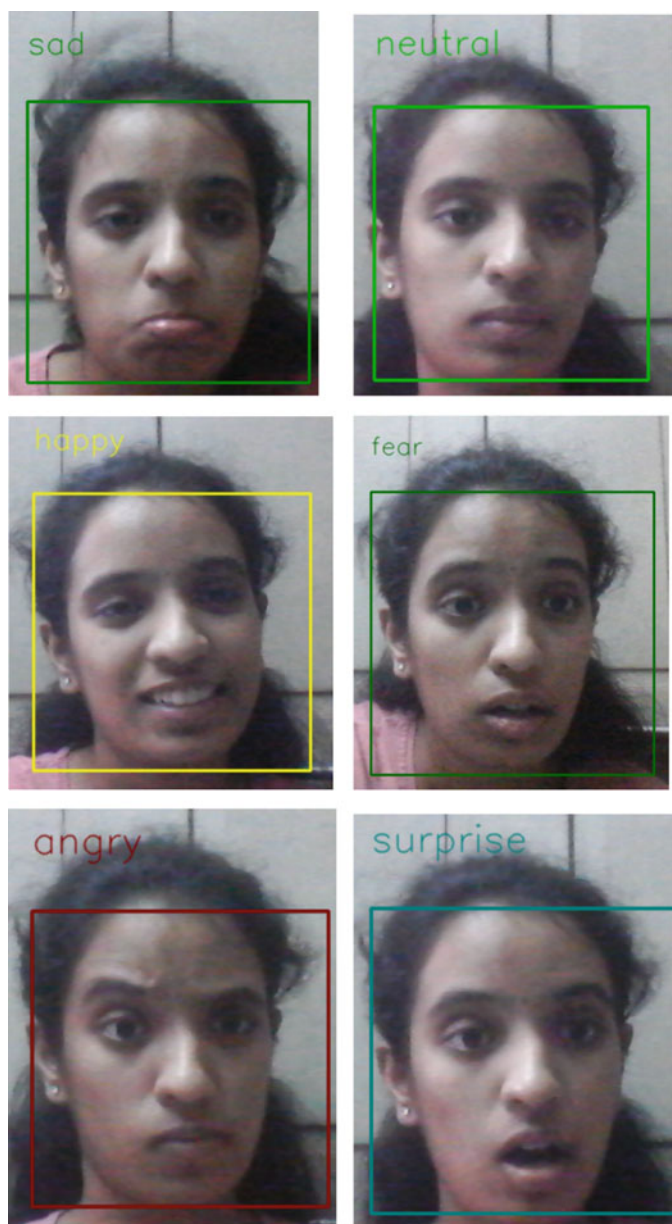
## 6 Conclusion

In this work, we combined gender classification and emotion detection using convolution neural network. The proposed architecture we included batch normalization and dropout. We also reduced number of parameters to be learned by using 1X1 kernel. The general CNN produced 95% accuracy on gender classification and 60% accuracy on emotion detection. The proposed model produced 96% accuracy on gender classification and 76% accuracy on emotion detection. This proposed method also reduced the learning parameters and training time. It outperformed the traditional machine learning algorithms.





**Fig. 7** Output of static image input showing the result “Neutral man, Happy and Neutral woman”



**Fig. 8** Real-time output captured by webcam showing “Angry,” “Fear,” “Happy,” “Neutral,” “Sad,” “Surprise” Emotions

## References

1. Arumugam D (2011) Emotion classification using facial expression. *IJACSA*) *Int J Adv Comput Sci Appl* 2(7)
2. Tsapatsoulis N, Karpouzis K, Stamou G, Piat F, Kollias S (2000) A fuzzy system for emotion classification based on the MPEG-4 Facial definition parameter set
3. Metallinou A, Katsamanis A, Wöllmer M, Eyben F, Schuller B, Narayanan S (2012) Context-Sensitive learning for enhanced audiovisual emotion classification
4. Bailenson\_ JN, Pontikakis ED, Mauss IB, Gross JJ, Jabon ME, Hutcherson CAC, Nass C, John O (2008) Real-time classification of evoked emotions using facial feature tracking and physiological responses, *Int J Human-Comput Stud* 66:303–317
5. Lien JJJ, Kanade T, Cohn JF, Li CC (2000) Detection, tracking and classification of action units in facial expression
6. <https://pathmind.com/wiki/neural-network>. Neural Network basics
7. <https://www.ijeat.org/wp-content/uploads/papers/v8i6S/F10190886S19.pdf>. Emotion recognition from facial expressions
8. Liliana DY (2019) Emotion recognition using neural networks. <https://iopscience.iop.org/article/10.1088/1742-6596/1193/1/012004/pdf>
9. Dhomne A, Kumar R, Bhan V (2018) Gender recognition through face using deep learning
10. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: *Proceedings of 2005 IEEE computer society conference on computer vision and pattern recognition, CVPR 2005 vol I*, no 8, pp. 886–893. <https://doi.org/10.1109/CVPR.2005.177>
11. Ajala A, Viriri S (1994) Face-based age and gender classification using deep learning model-olatunbosun, pre-conception
12. <https://towardsdatascience.com/paper-summary-a-theoretical-explanation-for-perplexing-behaviours-of-back-propagation>. NIPS2018/Google/Paper Summary
13. Rishil S (2020) Adaptive fuzzy network based transfer learning for image classification. In: *2020 IEEE international students' conference on electrical, electronics and computer science (SCEECS)*. IEEE
14. [https://opencvpythontutorials.readthedocs.io/en/latest/py\\_tutorials/py\\_objdetect/py\\_face\\_detection/py\\_face\\_detection.html](https://opencvpythontutorials.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html). Haar cascade details for beginners: developed by Paul Viola and Michael Jones back in 2001

# Blockchain-Based Authentication Scheme to Secure Supermarket Transactions—A Survey



L. Bilvashree, H. A. Brinda Nadig, K. R. Bharathi, B. S. Kavyashree, and K. Harshitha

**Abstract** Internet of things (IoT) takes an important part in authorizing in converting physical assets to smart establishments through the existing network infrastructures. IoT is aimed at preparing a multiplex information system along with sensor data acquisition, efficient data exchange through networking, artificial intelligence, big data, machine learning, clouds, etc. A system of storing information in a form that it can be protected from hacker system is blockchain technology. This technology stores the data in form of blocks and each block is connected to other block in a linear order. Sensor devices generate a huge amount of data which contains the private information such as transaction details and personal details. The collected information is stored in cloud, and it is essential to safeguard and secure this information. This paper gives a survey on various hashing algorithms and techniques used to provide security to data generated from an IoT device.

**Keywords** Blockchain · Cloud · Hashing · IoT · Security · Sensors

## 1 Introduction

Blockchain technology is a form of distributed ledger technology which has been gaining a lot of attention in areas beyond its cryptocurrency roots since many years: blockchain and Internet of things (*IoT*), blockchain and cybersecurity, blockchain and finance, and blockchain and logistics.

IoT is systemic network of physical devices or “things” implanted with sensors, softwares, electronics, and network connectivity that are able to accumulate and transmit data over a wireless network without human involvement. Objects can be sensed and controlled remotely in an extent across existing network infrastructure with the help of IOT, producing openings for undeviating unification between physical world and the computer-based systems, and hence getting improvement across efficiency, accuracy, and economic benefit as the result. Each object is uniquely perceptible via its implanted computing system but is also proficient to interact within

---

L. Bilvashree (✉) · H. A. Brinda Nadig · K. R. Bharathi · B. S. Kavyashree · K. Harshitha  
Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

actual foundation. Advanced connectivity of various devices and services that work beyond machine-to-machine communications (M2M) is offered by IoT which also provides coverage for a variety of protocols, domains, and applications.

In this paper, Sect. 2 shows the literature survey, and Sect. 3 shows comparative analysis of various authentication mechanisms and security systems for IoT environments.

## 2 Related Work

Latifi et al. [1] proposed a system to apply technique in commercial real estate. Compared to stock markets, real estates use almost twice the amount of investment. Holders, owners, and shareholders are hardly content with present process. The target was to check out trade of blockchain in real estate. Objective was to resolve simple problems and offer a lot more drives for a game theoretic table cost market. The number of shareholders in real estate much reduces because of the gaseous liquid and universal access. Here, they had implemented the concept of bitcoin digital cryptocurrency to build project. It includes endanger benefits with a stable charge coupon in real estate market was proposed that can be secured in case of manual activity of market illusion. This was possible to give liquidity to real estate market and get rid of agent simple issues in real estate. Operation coherence, network reliability and durability, rush settlement and bottom agreement values were other welfare. It was feasible to dispense foreign exchange to retail and put an end to intermediary—classical matters in plot. Blockchain automation advised to carry out this scheme to utilize the fashionable commitment.

Desai et al. [2] proposed a work on blockchain which had lately appeared to certain tool for a mixture of digital ledger construction for solitude-qualify and comprehensible deal, which can allow serious distributed applications without restoring the centralized trust. Individually put forward a story blend architecture, equivalently provides an isolation-intensify, liable sale where delicate proffer is unlocked on an independent, and popular cryptographic ledger are work for installments and responsibility. Public blockchain requires sanctioning many contrasting virtual currency. The sum of special irons and plain fetters troop can be legalized in case of ceremonial ballot and on condition that show up is a contradiction latter a seller statement will be discipline which will carry out the liability. Hybrid public transaction architecture cannot cross up to worn for extensive request, individually from transaction or separately from sales alike numerical capital exchange split secret known facts genuine flow fact documentation and registration.

Wang et al. [3] putforth a system for securing transactions. Blockchain technology has been used to represent the artchain. This platform has been built with registration, security and tracking, and provenance for artworks. This aims to provide a transparent transaction and identifying the artworks without leaving out the privacy and hence, this gives high-level architecture design, design of data model, and trust establishment. There are three layers in the first one; they are User Front End, followed

by ArtChain, and Trading Back End. Poa has been used to design the data model consisting of user, artwork and trading, and trust establishment. This model has been implemented using two types of tokenization—ACG2 and the utility token ACGT, also, upgradable smart contracts, initializing issue, safety control, privacy and confidentiality are some of the key features in this model. Eternal storage with proxy contract, separate logic and data in key–value pairs are smart contract methods that have been used in the project. Instead of the constructor, the concept of initializer function has been used.

Surianarayanan and Santhanam [4] proposed a work that concerns for the cloud security in an application that has been discussed, and also it aims to provide solutions for the same. Some of the problems dealt here are insecure application programming, data loss, and hijacking. The various security levels have been implemented for security and each level aims to provide required security at their level. Network level security is achieved using encryption techniques, firewalls, and network. The system level provides the required authentication and authorization at CSP end of the personnel's and at client side. Virtual machine provides security where isolation of applications and data is implemented for different clients. The layer called the application layer deals with proper mechanisms, i.e., applications which are loaded on virtual machines must be placed with proper mechanisms.

Li et al. [5] has discussed about the devices related to CA server which encounter single point failure attack, data of device interfere by inner attacks to address this issue blockchain technology to interfere safety ledger to IoT devices by allocating unique ID for each device to report blockchain. To design data protection mechanism in blockchain hashing significant was deployed. To verify proposed system implementation, prototype system based on hyper ledger fabric was used. This implementation includes environmental deployment, transaction on the chain, key generation technology based on CSPRNG and data integrity verification in blockchain performance analysis. Avert harmful nodes from intruder, withstand DDOS attacks, and prevent firmware backdoor advantageous were also implemented in proposed system.

Lee and Kim [6] the author states that information or device, validated data put onto a blockchain, data may be disclosed via repellent so that for a smart meter apply zero knowledge proof to prove without leaking input data such as public key and magnify concealment of blockchain for secrecy. In proposed system, the third party recognizes the address of user, and from this, intensity of power is utilized; also payment done is accessed by intruder so intend a system to safeguard the sensitive data adding a zero knowledge that certify data is accurate without providing data to verifier. Proposed a smart contracts and zero knowledge proof is used by system to secure the data, IOT is started on blockchain, which can avert authentication of IOT devices and prevent data tampering, zero knowledge proof is applied data over retrieval.

Deshpande, et al. [7] had proposed a work on sensors that are expended to perceive environment and system conditions. Analog signals offered by sensors are sent to android device. The sensors are used here for sensing the temperature, pressure, humidity, vibration, and intrusion. Admin sets threshold to each sensor implanted in the industry. The Android compares this threshold with the approaching the analog

signal. When it comes across turbulence in the condition, alert is spawned; and sends messages and email to admin. Later, adequate, and necessary actions are taken using artificial intelligence to solve the problems. This can only be made possible if there are former experiences or similar preceding conditions that are already stored in database. Here, cloud is the database used regarding its increased scalability.

Nagar and Alshamm [8] proposed a new concept called RSA-Key Generations Offline which has been used here and focuses to speed up RSA algorithm through an advanced method called new generation keys method. All the key values generated are stored in tables within database. Four security levels have been implemented, and each security levels has their own database and sets. E values and key length are used to identify these levels accordingly. RSA Handshake Database protocol has been used here. The role of this protocol is to create and update identical gateways database, level selections (Setid), and establish algorithm among gateways. Model has four steps—2 offline RSA-key generations, online encryption and decryption process. Two improved protocols have been proposed here, which are used to provide security against collusion attack. One of them is standard RSA key generation protocol, and another is unstable version. Two algorithms have been used here: modular reduction and modular multiplication. These algorithms are created on look-up tables. This enables arithmetic reckonings in byte-by-byte basis. A fresh concept called indexes exchange is used. This concept uses indexes exchange rather than using keys to exchange between different gateways.

Gondaliya and Savani [9] put forth the concept of using the products of more than two large prime numbers to decrease security attack on systems called the hybrid RSA—a new hybrid security cryptosystem. This has been proved more efficient than the earlier existing methods, i.e., RSA and HRSA. The parameters used in this model are key generation, encryption, and decryption time. These are used towards measuring efficiency. In this system, calculating and tracing value of  $M$  is tedious process, hence, providing more security than the previous methods, hence ensuring that this model provides confidentiality in the encryption and decryption process as shown in Fig. 9 [9]. This reduces direct attack and is more secure.

Zhang et al. [10] presented the idea of extended hash algorithm (EHA) combined with the Merkle Hash Tree for sheltering privacy. Extended hash algorithm is employed to vaguely certify and also in calculating the principle of the program. This is succeeded by a Merkle Hash Tree, where tree nodes conserve coherence, and tree values in non-leaf nodes are spawned by this algorithm by itself. When remote attestation is directed, just a node is acquired via encrypted hash value. For IMA architecture, a single inventory of “ $n$ ” number of nodes in total is needed to run the operation with time complexity of  $O(n)$ . Time complexity of extended hash algorithm combined with Merkle Hash Tree is  $O(\log 2n)$ . Complexity proves execution has maximum efficiency.

Eridani and Widiyanto [11] put forth the paper that deals with the monitoring system for sensors built using Raspberry Pi via MQTT protocol as a working implementation of IoT. MQTT protocol works by transporting the topic using publish and subscribe technique. The data is sent to the broker by employing publish procedure and the same is retrieved from the broker via subscribe mechanism. Raspberry Pi is the

**Table 1** Comparative analysis of different authentication mechanisms

S. No.	Title of the paper	Methodology	Limitations
1	ArtChain: blockchain enabled platform for art marketplace [3]	The Hyperledger Fabric is used to implement the system. They have designed and implemented a utility token called ACGT to achieve high performance	Anti-counterfeiting for original works of art by integrating with smart modules of IoT and activating relevant smart hardware and other functionality by artists or collectors is not provided
2	Security issues and control mechanisms in cloud [4]	Various security levels have been implemented to achieve this	Even though this is secure more measures with respect to security must be implemented
3	A blockchain—based authentication and security mechanism for IOT [5]	Design a data protection mechanism by hashing significant data into a blockchain	Scalability issue relating to size of blockchain ledger that might lead to centralization
4	Implementation of IOT system using blockchain with authentication and data protection [6]	Apply a zero-knowledge proof to smart met	Processing of power and time required to perform encryption algorithm objects involved in blockchain is more
5	Industrial automation using Internet of Things (IOT) [7]	Analog signal is provided to android device produced by sensors	Wi-Fi/Internet connections fluctuation may create problems. SMS/Email Alerts may have range problem

chief controller here. This sector is linked to various sensors; two sensors in this situation: IR and temperature. Both sensors are interpreted in voltage. Hardware is then calibrated to transfer data via MQTT protocol to cloud server. Broker in computer can be used to prefect televised data. HiveMQ is broker here (Table 1).

### 3 Conclusion

Based on the literature survey carried out and ongoing state of network security, we can deduce there is a scope for working on the security of network. The obligation here is to build a secure connection between the IOT environment and the users to interact via IoT gateway managed by admin.

The system which is going to be proposed is expected to secure transactions by using Adler-32 hashing algorithm. The hash values generated are cross verified with primary hash values to make sure the transaction details are not modified. This system will be very difficult for hackers to meddle with as cloud and blockchain are also incorporated into the system design. In case the details have been modified by the third party, the original data can be retrieved



from the copy stored in the cloud server. This system is expected to be cost effective and give fast execution time

## References

1. Latifa S, Zhang Y, Cheng LC (2019) Blockchain based real estate market: one method for applying blockchain technology in commercial real estate market. In: IEEE international conference on blockchain
2. Desai H (2019) Murat anteriorly and Lallana Kagal: a hybrid blockchain architecture for privacy-enabled and accountable auctions. In: IEEE international conference on blockchain
3. Wang Z, Yang L, Wang Q (2019) ArtChain: blockchain enabled platform for art marketplace. In: IEEE international conference on blockchain
4. Suba Surianarayanan, T.Santhanam (2012) Security issues and control mechanisms in cloud. In: International conference on cloud computing technologies, applications and management
5. Li D, Peng W (2018) A Blockchain-based authentication and security mechanism for IoT. In: 27th International conference on computer communication and networks
6. Lee CH, Kim KH (2018) Implementation of IoT system using block chain with authentication and data protection. In: International conference on information networking
7. Deshpande A, Pitale P, Sanap S (2016) Industrial automation using Internet of Things (IOT). Int J Adv Res Comput Eng Technol
8. Nagar SA, Alshamma S (2018) High speed implementation of RSA algorithm with modified keys exchange. In: 1st International conference on data intelligence and security
9. Gondaliya J, Savani J, Dhaduvai VS, Hossain G (2018) Hybrid security RSA algorithm in application of web service. In: 1st International conference on data intelligence and security
10. Zhang Y, You Y, Yi I, Wang L (2017) A remote-attestation-based extended hash algorithm for privacy protection. In: International conference on computer network, electronic and automation
11. Eridani D, Widiyanto ED (2018) Performance of sensors monitoring system using raspberry Pi through MQTT protocol. In: International seminar on research of information technology and intelligent system

# Unified Flow Based Approach in Software-Defined Networks



V. Deeban Chakravarthy, Dara Sujith Chandra,  
and Seemakurthi Sri Sathya Pavan

**Abstract** We investigated energy-efficient strategies for transferring data packets between network devices and access points in order to reduce transfer latency and extend the network device's energy competency. It is demonstrated that the resource distribution challenge for network methods under the corrected routing approach is convex. On this starting point, a decisive vector algorithm is constructed, whose implementation is equivalent to the finest possible result for comprehensive routing protocol and aspires to mutually enhance the energy efficient routing policy and network device allocation. The difficulty of the existing system is evaluated and geared towards a suitable architectural methodology. Mininet Emulator output confirm the convergence performance and energy efficient functionality of the proposed decisive vector optimization algorithm.

**Keywords** Energy-efficient software-defined networks · Server provisioning

## 1 Introduction

The Internet has become a significant piece of regular daily existence. From looking through Google to purchasing from Amazon to staying aware of companions on Facebook, it is a piece of individuals' everyday lives. Google has even gotten productive enough to warrant turning into a word reference term. Behind every one of these multibillion dollar organizations are networks that need to adjust to always evolving needs. Since the Internet turns out to be more priceless than any other time in recent memory, it is significant for it to be more available and simpler to keep up.

---

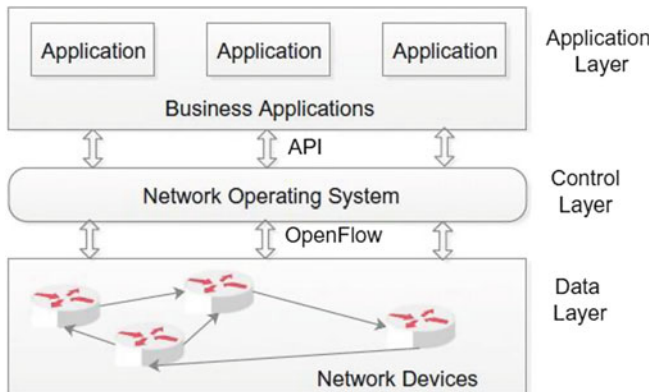
V. D. Chakravarthy (✉) · D. S. Chandra · S. S. S. Pavan  
SRM University, Kattankulathur, Chennai, Tamilnadu, India  
e-mail: [deepanv@srmist.edu.in](mailto:deepanv@srmist.edu.in)

D. S. Chandra  
e-mail: [dd1010@srmist.edu.in](mailto:dd1010@srmist.edu.in)

S. S. S. Pavan  
e-mail: [ss5800@srmist.edu.in](mailto:ss5800@srmist.edu.in)

The networks knows the foundation of the Internet, need to adjust to changes, without being massively work concentrated in equipment or software changes. By refreshing switches with the goal that the data plane and the control plane can be isolated, a unified controller can make more advanced courses to fittingly advance traffic. These forwarding tables can utilize rule contribution to the controller to advance routing and make networks more proficient.

Generally, the data and control plane components were limited in one network component (switch/router) and vendor specific. A particularly shut construction of the framework empowers the entrance just to the merchants, to change the network as indicated by client prerequisites. SDN conquers these obstacles and offers a methodology where the control plane and the data plane are isolated giving a focal control point (called controller) for halfway organizing and dealing with a network. The communication between the control plane and the data plane is finished utilizing a south bound application programming interface (API) protocol like the OpenFlow protocol.



Software-defined networking vows to significantly improve on network control, management and empower advancement through network programmability. PC networks are normally developed from an enormous number of network gadgets (like switches, routers and firewalls) with numerous intricate protocols (software), which are carried out and installed on them. Network engineers are liable for designing approaches to react to a wide scope of network occasions and application situations. They physically change these undeniable level arrangements into low-level design orders. These perplexing undertakings are frequently cultivated with admittance to extremely restricted devices. In this manner, network management control and execution tuning are very difficult and mistake inclined undertakings.

Forwarding of a data parcel is cultivated by a forwarding gadget (switch/router) situated in the data plane and the controller situated in the control plane is answerable for taking a choice about how the bundle ought to be directed in the network and push this choice to the forwarding gadget.

## 2 Routing Protocol

### 2.1 Routing Information Protocol

It is a similarly straightforward routing protocol that was broadly sent in little networks during the 1990s. Every router in the routing protocol-controlled routing domain intermittently communicates its whole routing table on the entirety of its interfaces. These transmissions incorporate the bounce check from the telecom router to each reachable network. The heaviness of each bounce can be customized to oblige important contrasts between the jumps, for example, interface speed. Adjoining routers can incorporate this reachability and jump tally information into their own routing tables, which will thus be spread to their neighbours. This proliferation invades the whole routing domain or self-governing framework. The routers in an AS offer nitty gritty routing information, permitting every part router in the self-sufficient framework to process the most limited way to any objective inside that AS. In a stable network, in the long run, all the routing tables in that routing domain will unite, and the bounce tally information can be utilized to decide a most minimal expense course to arrive at any network in that domain. Each router uses the weighted bounce to examine the distance from more than one interface to the objective network and selects the best next-jump router as one that provides the most out distance to the objective. Routing protocol does not have to have a total image of the network geography to make this distance-vector routing calculation. The MinMax routing response time is defined by

$$\delta_{i,j} = \varphi_{i,j} + \varphi_{i,j} \frac{\rho_{i,j}}{1 - \rho_{i,j}}$$

The median user degree is defined as

$$T_k = \sum_{j=1}^M \sum_{i=1}^N r_{k,i} w_{i,j} v_{i,j}^k u_{i,j}^k / \sum_{i=1}^N r_{k,i}$$

Undoubtedly, since these control protocols are really distributed protocols, when something is purposefully changed, for example, a worker being added or a communications connect removed from administration, these distributed protocols set aside some effort to meet to a bunch of reliable forwarding tables, unleashing transitory devastation inside the data place. In this manner, we arrive at a circumstance where most geography change is done automatically and purposefully, lessening the advantage of the self-ruling and distributed protocols, and the size of the networks is bigger than these protocols were intended for, making intermingling times inadmissibly long.

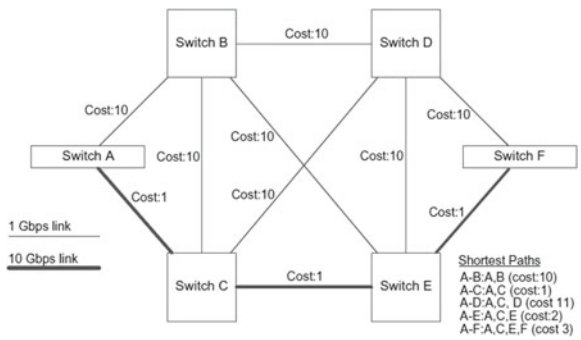
2.2 *Spanning Tree Protocol STP*

Essential layer two forwarding, otherwise called transparent bridging, can be performed freely by each switch in the network. Nonetheless, certain geographies require a burden of a chain of command on the network to forestall circles, which would cause broadcast radiation. The spanning tree protocol is an illustration of the activity of self-sufficient gadgets taking an interest in a distributed dynamic interaction to make and authorize a chain of importance on the network. The outcome is the right activity of straightforward crossing over all through the domain to the detriment of union inertness and potentially discretionary arrangement. This arrangement was a trade-off among cost and intricacy.

Various ways might have been upheld yet at more noteworthy expense. Spanning tree protocol was sufficient when networks were of more limited size, yet as networks became the crossing tree arrangement got dangerous. These issues show themselves in a striking design when networks arrive at the size of the advanced data community. For instance, IEEE 802.1D determines the accompanying default clocks for STP: 15 s for tuning in, 15 s for learning and 20 s for max-age break. In more seasoned networks, combination seasons of 30–50 s were normal. Such postponements are not acceptable in the present data habitats. As the size of the layer-two network develops, the probability of more noteworthy postpones increments.

2.3 *Shortest Path Calculation*

Shortest path calculation achieves this objective to build a chart addressing the layer two connection state geography. When this diagram exists, most limited way computations are clear, however, more perplexing than with crossing tree.



To further explain what we mean by most brief way counts, we illustrate a basic chart in figure that can be used to determine the much more concise way in a network with five switches. The costs allotted to the different connections might be allocated

their qualities as per various standards. A basic model is to make the expense of a network interface conversely relative to its transfer speed. Consequently, the expense of traveling a 10 Gbps interface is one-tenth that of traveling a 1 Gbps connect. At the point when the most limited way estimation is finished, the hub playing out the figuring realizes the smallest expense way to any of different hubs in the network. The most minimal expense way is viewed as the briefest way. For clearness, we should call attention to that is utilized in the SPB setting carefully for layer two way count. This contrasts from its traditional application in computing layer three courses

### 3 Related Work

The main concept for software-defined network was derived from the Ethane of Project Clean Slate [3] Program at Stanford University which puts control into the hands of users, owners and applications.

The objective of Ethane is to assemble another engineering for big business networks to give an incredible and straightforward management model [5, 8] and solid security ensures. In this undertaking, Martin Casado developed assembling one brought together SDN by which network chairmen can tailor stream-based security approaches to their individual necessities [3, 16]. As a result of these considerations, we believe that a critical approach is required in software-defined networks.

The data plane gives rich controller programming models and deliberations to oversee equipment assets. OpenFlow is normally seen as one general deliberation of the data plane that gives standard interfaces to introducing and erasing rules [6, 9] in the stream table. Notwithstanding, there are a few downsides of OpenFlow. To beat these disadvantages, different deliberations of the data plane have been proposed. As an alteration of OpenFlow, an alternative flow controller [5, 11, 14] was proposed to lessen pointless expenses to address the issues of superior networks. Since the deliberation of calculation and capacity assets are disregarded in OpenFlow, label cast protocol was proposed to all the more likely help future networks [1] by abstracting forwarding assets just as calculation and capacity assets in the data plane. Various working group proposed the forwarding and control element separation protocol [7] to accomplish similar targets as programmable network protocol. Associated with OpenFlow, a programmable network protocol, forwarding and control element separation protocol [5, 7, 13, 16] have a unique model that makes its protocol very amazing; in any case, it needs open source accessibility [14] for experimentation. Since OpenFlow-based devices are costly, it is unthinkable for organizations to supplant their current network gadgets with OpenFlow-based network devices.

As a result of the creation of OpenFlow network assessment devices to recognise the utility of management [6]. Y.W. Wang, Gong, et al. applied controller in the transporter grade network, stretching out the controller engineering to help multi-supplier network management [6, 10, 14] capacities in transporter networks. Kim et al. introduced an occasion driven network control structure, specifically Lithium, which makes network management simpler [11]. Fernandez assembled a proving

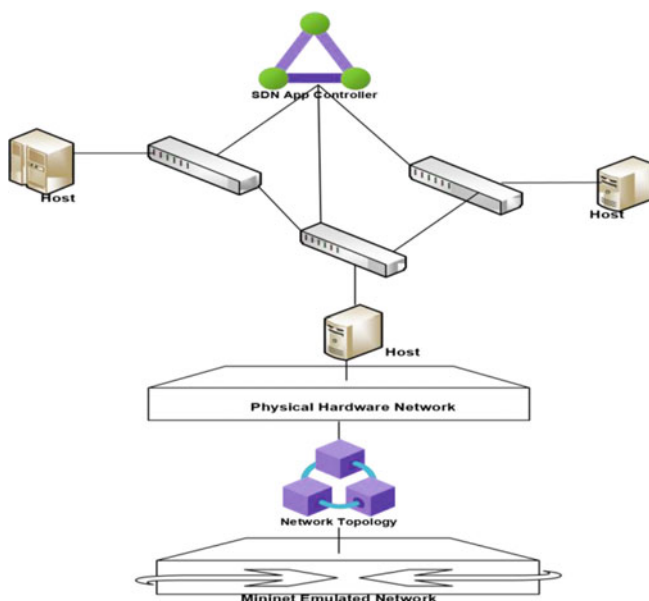
ground for conveying estimations and applications in broadband access networks. This proving ground can assemble information about network topology [2, 6, 14], accessibility, reachability, etc. Kim et al. planned and carried out to work on different parts of network activities and management and fills [12] in as the paste between significant level network strategies and low-level network designs.

L. F. Muller, R. R. Oliveira, M. C. Luizelli, et al. applied controller to the advancement of Internet of things applications [7]. Lange, S. Gbert, Xie, D. Guo, Hu et al. assembled a portable cloud management framework [9] based on SDN, which profits by the powerful setup of OpenFlow switches. Wang et al. tended to the issues of versatility [7] in Internet protocol (IP) networks utilizing SDN while planning an OpenFlow-based portability protocol. Das et al. introduced an SDN-based network management system named Flow for large data handling. M. F. Bari, A. R. Roy, S. R. Chowdhury et al. incorporated a network control work given by a SDN controller into Hadoop [18, 19] to together streamline execution of huge data preparing and network use.

## 4 Details of Proposed Operations

Based on the accompanying attributes, a distance-vector algorithm is used to present the results:

1. Distance—How far the objective network is based upon a metric, for example, bounce check.
2. Vector—The course (next-jump router or departure interface) needed to get to the objective. This routing information is traded between straightforwardly associated neighbours



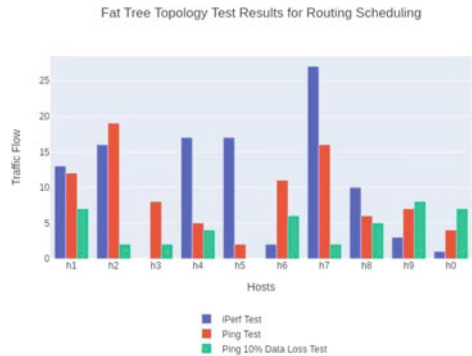
Therefore, when a hub gets a routing update, it has no knowledge of where the neighbour took in it from. At the end of the day, the hub has zero ability to see of the network past its own neighbour.

## 4.1 Shortest-Path Bridging

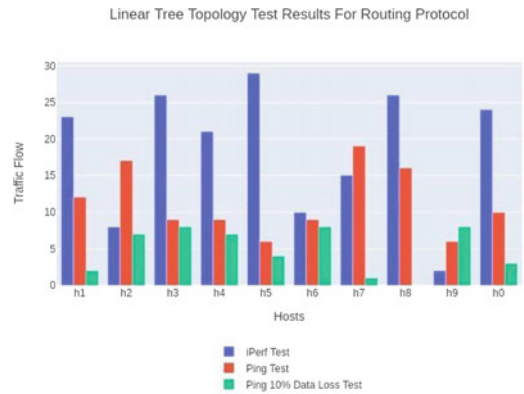
STP permitted just a single dynamic way to an objective, experienced moderately lethargic assembly times and was limited to little network geographies. Albeit the fresher executions of STP have improved the combination times, the single dynamic way inadequacy has been tended to in another layer two protocol, shortest-path Bridging is a system for permitting numerous simultaneous ways through a two-layer texture by means of communitarian and distributed figuring of briefest and most productive ways at that point dividing that information between the partaking hubs in the fit network. This trademark is called multipath.



5 Experiment Results



The above experimental result shows the fat-free topology test results for routing scheduling. The above outcome is easy to analysis. In the above scenario, in Host h1 the iperf managed to perform data transfer 12.5 Giga bytes when there is a normal traffic. In Host h7, the iperf is managed to transfer 22.4 Giga bytes when there is no minimum traffic in the network. In Host h0, the iperf managed to transfer 0.97 Giga bytes when there is a heavy traffic in the network.



The above experimental result shows the linear tree topology test results for routing scheduling. The above outcome is easy to analysis. In the above scenario, in Host h1 the iperf managed to perform data transfer 226 Giga bytes when there is a normal traffic. In Host h5, the iperf is managed to transfer 28.2 Giga bytes when there is no minimum traffic in the network. In Host h9, the iperf managed to transfer 1.89 Giga bytes when there is a heavy traffic in the network.

In the above two topologies, linear tree topology is best suited for this experimental scenario. For all performance test considered, the traffic flow improves when

adjusting the linear tree topology form, after which it seems that improving the traffic flow size does have the identical substantial effect.

## 6 Conclusion

The comprehensive literature component of this study critically highlights a wide range of network approaches and their functions for constructing energy of existing issues. Assuming that the network can support the service, delivery can be activated. During service operation, two monitoring and control loops run concurrently: one at network level and the other at service level.

In this research, routing and energy-efficient control procedures are implemented directly on emulated network devices using a range of automated and manual methods.

## References

1. Feamster N, Rexford J, Zegura E (2014) The road to SDN: an intellectual history of programmable networks. In: ACM SIGCOMM computer communication review, vol 44, no 2, pp 87–98
2. Lara A, Kolasani A, Ramamurthy B (2014) Network innovation using openflow: a survey. *IEEE Commun Surv Tutor* 16(1):493–512
3. Ethane: a security management architecture, a stanford clean slate project. <http://yuba.stanford.edu/ethane/>
4. Kreutz D, Ramos FMV, Verissimo PE et al (2015) Software-Defined networking: a comprehensive survey. In: *Proceedings of the IEEE*, vol 103 no 1, pp 14–76
5. “Giroire F, Moulhierac J, Phan TK (2014) Optimizing rule placement in software-defined networks for energy aware routing. In: *Proceedings of the 2014 global communications conference (GLOBECOM)*, IEEE, pp 2523–2529
6. Shah SA, Faiz J, Farooq M et al (2018) An architectural evaluation of SDN controllers. In: *The 2013 international conference on communications (ICC)*. IEEE, pp 3504–3508
7. Fernandez MP (2018) Comparing Openflow controller paradigms scalability: reactive and proactive. In: *The 27th international conference on advanced information networking and applications (AINA)*. IEEE, pp 1009–1016
8. Monaco M, Michel O, Keller E (2018) Applying operating system principles to SDN controller design. In: *Proceedings of the 12th workshop on hot topics in networks*. ACM
9. Xie J, Guo D, Hu Z et al (2017) Control plane of software defined networks: a survey. *Comput Commun* 67:1–10
10. Binh LH, Duong TTV (2021) Load balancing routing under constraints of quality of transmission in mesh wireless network based on software defined networking

# Towards a Framework for Context-Aware Intelligent Transportation System: Case of Kigali



Gatera Antoine , Chomora Mikeka , Gaurav Bajpai , Andras Valko, and Kayalvizhi Jayavel 

**Abstract** The intelligent transportation system (ITS) proved to be a cost-effective transportation solution. Artificial intelligence (AI), machine learning (ML) and the Internet of things (IoT) are supposed to take advantage of big data's accessibility for successful data processing. Development of platforms compatible with innovative services will improve living standards. Research in the ITS field together with their deployment is focused on developed countries. It is important to explore the concept in developing countries. In this paper, a survey was conducted using a random sampling technique to obtain opinions, readiness for ITS, existing technologies and transportation applications. The analysis of the results showed that 73.78% are not familiar with ITS. An overview of the challenges for intelligent transport in the region with a specific focus on road safety is presented. An IoT-based conceptual framework is proposed to improve the transportation system specifically in Rwanda. This research contributes to raising awareness of the possibility of ITS.

**Keywords** ITS framework · IoT · Quality of service

---

G. Antoine (✉) · A. Valko

African Center of Excellence in Internet of Things, College of Science and Technology,  
University of Rwanda, Kigali, Rwanda

C. Mikeka

Department of Physics, University of Malawi, Zomba, Malawi

G. Bajpai

Department of Computer and Software Engineering, College of Science and Technology,  
University of Rwanda, Kigali, Rwanda  
e-mail: [gaurav.bajpai.2014@ieee.org](mailto:gaurav.bajpai.2014@ieee.org)

K. Jayavel

Department of Information Technology, SRM Institute of Science and Technology, Chennai  
603203, India  
e-mail: [kayalvij@srmist.edu.in](mailto:kayalvij@srmist.edu.in)

## 1 Introduction

Urban population growth in Africa is projected to reach 60% by 2050 [1]. Urban projection of the Rwanda's working-age population is expected to increase from 1.6 million (2019) to 3.2 million in (2032) [2]. In Rwanda's perspective, a smart city is "a compact, connected, socially inclusive and resilient city that leverages the power of technology, data and innovation to improve the quality of life of residents [3]. A smart city to attain, and there is a need for intelligent systems. In this paper, a survey was conducted to understand attitudes towards the applicability of some ITSs in Rwanda. Discussion of the findings highlights good ways for the country to develop ICT implementation strategies for intelligent systems. A contextual ITS framework that assists in creating a set of activities required to provide real-time traffic information is proposed".

This paper is organized as follows: Sect. 2 presents a review of ITS-based applications and technologies. Section 3 presents the road transportation and safety, and Sect. 4 proposes an IoT-based framework for ITS. Section 5 presents the methodology used, Sect. 6 presents the analysis of the survey results, and finally Sect. 7 presents the conclusion.

## 2 Review of ITS Applications and Technologies

In transportation point of view, there is a requirement for roadside units (RSUs) capable of communicating with the vehicles [4]. Autonomous vehicles (AV) are expected to reduce crashes, pollution and congestion [5]. The use of electric cars will reduce the environmental pollution that was due to CO<sub>2</sub> emission [6]. Transport contributed about 23% of all climate emissions [7]. Vehicle to everything (V2X) communication ensured easy mobility [8]. The African Union (AU) calls on African states to invest in infrastructure for innovation with technical improvements. Intelligent in-vehicle and roadside traffic signals used adaptive traffic light control agents to produce appropriate control signals [9]. A way to predict bus arrival time is important for reducing wait times at bus stops [10]. An in-depth learning approach can improve the accuracy of traffic forecasts to mitigate traffic congestion [11]. Internet of the road has profound benefits such as accident prevention and time saving [12] and automatic route crack detection methods for maintenance of the road [13]. The Global Positioning System (GPS) to monitor track speeding vehicles in real time will help in identifying safe driving drivers [14]. The IoT system can easily be developed based on a Raspberry Pi, control cameras, infrared (IR) sensors and radio frequency identification (RFID) technologies [15]. The RFID tags are used to prioritize special emergency vehicles such as ambulances [16]. Future mobility in Rwanda will depend on (i) efficient road network, (ii) regular public transport availability with affordable services and (iii) effective institution's responsibility for urban transport planning and policy-making.

### 3 Road Transport and Safety in Sub-Saharan Africa

Road safety causes a substantial cost to economies together with great trauma to societies. The average cost of road traffic accidents in developing countries is estimated at 3% up to 5% of the gross national product. The safety systems applications are the core of the automated driving technology [17]. Road safety depends on driver behaviour, infrastructure quality and regulatory framework [18]. Improvement in road infrastructure is critical to overall road safety improvement [19]. Countries need to align policies and framework plans for the transportation sector [20]. Transportation has to be conceived from a broad point together with other relevant aspects related to sustainable development [21].

#### 3.1 Transport Sector in Rwanda

In Rwanda, transport is operated in three modes, namely (i) land transport, (ii) air transport and (iii) maritime transport. The Government of Rwanda (GoR) is conscious of the importance of transport infrastructure development in delivering on its long-term development vision. The transportation sector keeps facing challenges, including but not limited to (i) public transport that is characterized by delays, inaccessibility, unpredictability, (ii) lack of streamlined bus schedules, delays at terminals, fares, passenger information and (iii) lack of integrated public transport [22]. Rwanda's transport safety is essential due to which motorized and non-motorized traffic share the same space, although they are different in operational speed and knowledge of traffic regulations. Table 1 presents the statistical ratios of accident cases for in the past four years.

**Table 1** Road traffic accidents in Rwanda (2016–2019)

Categories	2016	2017	2018	2019	Mean	Std.Dev	Variance
Fatal	593	549	597	673	603	51.4	2650
Serious injury	629	733	885	911	789.5	132.7	17611
Minor injury	1794	1896	1887	1485	1765.5	192.6	37095
Properties damage	3347	2577	2242	1584	2437.5	773.3	537777

### 3.2 *Existing Policies and Preferences*

The GoR has separated functions by allocating policy, implementation, regulatory responsibilities to a number of agencies, some being highlighted. They are grouped based on their functional line as follows: (i) political layer consisting of The Ministry of Infrastructure (MININFRA) responsible for formulating integrated transport and corresponding strategies; (ii) authority layer consisting of Rwanda Transportation and Development Agency (RTDA) in collaboration with the Utilities Regulatory Authority (RURA). RTDA performs all tactical functions for public transport, (iii) infrastructure development layer by RTDA and City of Kigali and (iv) enforcement layer by the Rwanda National Police (RNP) for safety and security [23]. The Rwanda ICT Strategic and Action Plan (NICI III-2015) shows that leverage ICTs in all sectors will increase private sector productivity [24]. Therefore, identifying opportunities that enable a sustainability is recommended [3]. The Rwanda market for ICT is characterized by a high evolution and demand for services. 4G LTE technology is deployed on 94.2% of geographic coverage and 96.6% of population coverage. Enhancement of ICT integration in public transport, accidents and real-time information is needed.

## 4 **Proposed Conceptual Framework for ITS-Based IoT**

Transport modelling together with traffic forecasting is the backbone of road planning. Technology adoption in transportation is always challenging for technologists, managers and executives. ITS applications discussed in Sect. 2 are designed to improve the efficiency and safety together with the sustainability of road transport. Smart infrastructure must be considered in road planning and design to enable the road's communication capabilities. Behind intelligent infrastructure, the ITS framework at the country level has to be developed to assist road authorities to monitor road traffic. Figure 1 depicts a comprehensive map to understand key factors to be considered in developing a transportation master plan. In this paper, the IoT-based ITS conceptual framework presented provides reasonable tools to understand the situation in sub-Saharan cities.

### 4.1 *Input Layer*

This is the data ingest layer. Integration of the IoT with road traffic information is collected through the use of various technologies (inductive loop detectors, radar detectors, magnetometers, weight in motion systems and speed detectors) to provide the status of road network issues. The entire traffic data source is gathered for data value harvesting and extraction. Public transit agencies and private sector tech firms

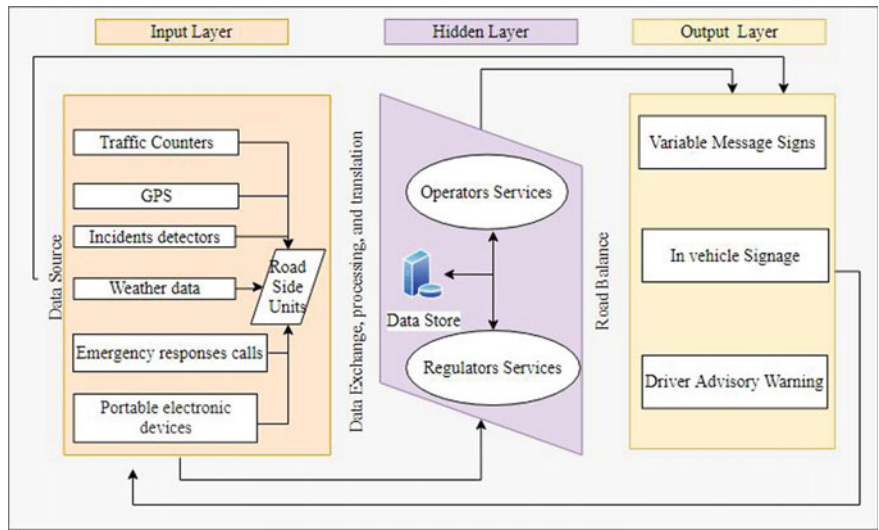


Fig. 1 Conceptual framework of an ITS-based IoT

can use the technology to simplify passengers to plan for their trip, but also the government to plan for the future based on this data.

4.2 Hidden Layer

Traffic data is generated in various forms. The data is gathered for further translations at this processing layer. An IT system is built to distinguish pieces of information together with enabled software for a particular purpose. Data transmission to disseminate processed data is performed at this layer. All road users, including passengers, car drivers, cyclists across all modes of road transport would benefit from the effective use of generated data.

4.3 Output Layer

The data acquired after being processed is shared according to the needs of the user of the route. Access to information through devices is possible with Wi-Fi, LTE communication access for revenues, roadside variable messages, social media messages and in-vehicle displays. Today, new technologies paired with existing data sources give the world an unprecedented opportunity to address the increased demand for data. The framework proposed is based on the survey results presented in Sect. 6.

## 5 Data and Method

To conceptualize the ITS framework for smart city, in-depth study was used through government documents, interviews and the survey in Kigali city. The survey aimed to understand attitudes towards ITS and the level of ITS concept awareness in Rwanda. The population of interest included professionals from the government, private sector, academia and the public sector. The government agencies whose primary responsibilities are related to transportation strategy, policy, implementation and service delivery were chosen as they have the ability to influence traffic management practice.

## 6 Results and Discussions

The survey responses were analysed to find issues and challenges to be addressed for ITS implementation in Rwanda. The respondent's perceptions were assumed to be corresponding to their critical perception of ITS. Out of 400 responses, 225 responses were suitable for the analysis. 28% were from the government, 28% from the industry/private sector, 23.11% from academia and 20.89% from the public sector. The analysis of respondents responses was analysed by checking the collinearity of the variables and respondents awareness towards them. The hypothesis was that there is a strong relationship between ITS and the reviewed technologies. Figure 2 shows Pearson's product-moment correlation ( $r$ ) for selected experimental variables. The full correlated data has a value of 1, while the worse correlated data has around zero value. Few parameters are correlated in the sense that gives an insight into their relationship to ITS. The results show that correlation coefficients tend to zero instead of 1.

ITS implementation is a gradual process, and there is a dependent phenomenon that includes the context awareness of ITS system and applications. The data presents challenges to be addressed in implementing the ITS. A typical context-awareness system is built based on three layers: the perception, the reasoning layer and the acting layer. Of the 220 respondents who preferred to share their education level, 1.4% were with primary, 18.6% secondary and 80% university. To determine respondent's familiarity with ITS, out of the 225 respondents, 26.22% were familiar and 73.78% were not familiar to ITS. To examine the respondent's awareness level about the technologies discussed, a five (5) point Likert-type scale was used. Table 2 presents the level of the respondent's familiarity with the technologies respondents.

However, there is little chance that a standard ITS will succeed in sub-Saharan Africa. ITS has a global dimension; sub-Saharan countries do not meet ITS standards for wireless communications, network architecture and protocols. The creation of a common agreed standard for the data collection, control of vehicle networks, processing and provision of information would harmonize ITS. Infrastructure and their management have been a challenge over the years for the least developed countries.



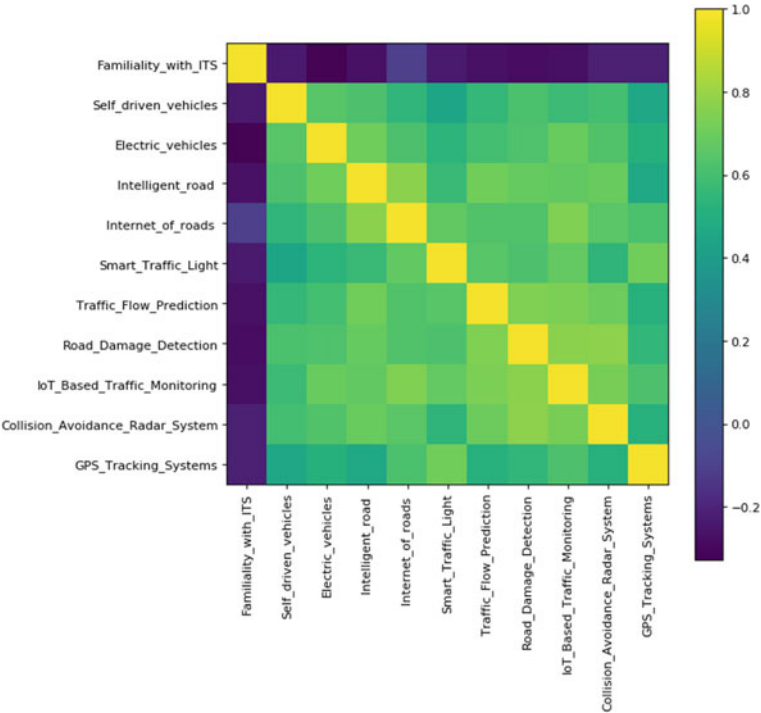


Fig. 2 Pearson correlation between variables

Table 2 Technologies and systems awareness

Technology	Aware (%)	Not aware (%)
GPS tracking systems	51.1	48.9
Smart traffic lights	49.3	50.7
Self-driven vehicles	38.7	61.3
Internet of roads	37.3	62.7
IoT-based traffic monitoring	35.6	64.4
Traffic flows prediction	33.3	66.7
Electric vehicles	32.4	67.6
Intelligent road (RSUs)	29.8	70.2
Road damage detection	27.4	71.6
Collision avoidance system	26.7	73.3

Various goals and initiatives fail due to which governments try to tackle each complexity in one solution. The realization of ITS is beyond a political will. There is a need to combine backgrounds and domains to develop and build technology-friendly infrastructures.

## 7 Conclusion

Technologies are expected to bring major changes to road safety. This paper explored the process for achieving intelligent transport. Practitioners' and decision-makers' perceptions of ITS were assessed, since they are ITS influencers. The IoT-based ITS platform was proposed to improve transport. The results presented an interesting perspective of training industry professionals and decision-makers about the technical requirements for intelligent mobility. This research will assist municipal managers and program directors to better understand the complex concept of intelligent transportation by breaking down the process into process objectives, components and enablers. Smart city indicators include smart mobility, smart people, smart governance and a smart environment.

**Acknowledgements** This study was funded by the African Center of Excellence in Internet of Things established at College of Science and Technology, University of Rwanda.

## References

1. United Nations. World Urbanization Prospects The 2014 Revision-Methodology, Online available: <https://population.un.org/wup/Publications/Files/WUP2014-Methodology.pdf>. Accessed on 20 April 2021
2. Jonathan B, Sally M, Robert B, Laura W (2021) Housing need in Kigali - C-38406-RWA-1. Available online: <https://theigc.org/wp-content/uploads/2019/07/Bower-et-al-2019-Final-report.pdf>. Accessed on 2 Mar 2021
3. Smart City Rwanda Masterplan. Online available: [https://unhabitat.org/sites/default/files/documents/2019-05/rwanda\\_smart\\_city-master\\_plan.pdf](https://unhabitat.org/sites/default/files/documents/2019-05/rwanda_smart_city-master_plan.pdf) Accessed on 15 Jan 2021
4. Bagloee SA, Tavana M, Asadi M, Oliver T (2016) Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *J Modern Transp* 24(4): 284–303. <https://doi.org/10.1007/s40534-016-0117-3>
5. Fagnant DJ, Kockelman K (2015) Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp Res Part A: Policy Pract* 77: 167–181. <https://doi.org/10.1016/j.tra.2015.04.003>
6. Habib S, Khan MM, Abbas F, Sang L, Shahid MU, Tang H (2018) A comprehensive study of implemented international standards, technical challenges, impacts and prospects for electric vehicles. *IEEE Access* 6: 13866–13890. <https://doi.org/10.1109/ACCESS.2018.2812303>
7. Tintelecan A, Dobra AC, Marțiș C (2019) LCA indicators in electric vehicles environmental impact assessment. In: 2019 Electric Vehicles International Conference (EV), pp 1–5. <https://doi.org/10.1109/EV.2019.8892893>

8. Huang W, Li P, Zhang T (2018) RSUs placement based on vehicular social mobility in VANETs. In: 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp 1255–1260. IEEE. <https://doi.org/10.1109/ICIEA.2018.8397902>
9. Mousavi SS, Schukat M, Howley E (2017) Traffic light control using deep policy-gradient and value-function-based reinforcement learning. *IET Intelligent Transp Syst* 11(7): 417–423. <https://doi.org/10.1049/iet-its.2017.0153>
10. Pang J, Huang J, Du Y, Yu H, Huang Q, Yin B (2018) Learning to predict bus arrival time from heterogeneous measurements via recurrent neural network. *IEEE Trans Intelligent Transp Syst* 20(9): 3283–3293. <https://doi.org/10.1109/TITS.2018.2873747>
11. Lv Y, Duan Y, Kang W, Li Z, Wang FY (2014) Traffic flow prediction with big data: a deep learning approach. *IEEE Trans Intelligent Transp Syst* 16(2):865–873. <https://doi.org/10.1109/TITS.2014.2345663>
12. Lozano Dominguez JM, Mateo Sanguino TJ (2019) Review on V2X, I2X, and P2X communications and their applications: a comprehensive analysis over time. *Sensors* 19(12):2756. <https://doi.org/10.3390/s19122756>
13. Shi Y, Cui L, Qi Z, Meng F, Chen Z (2016) Automatic road crack detection using random structured forests. *IEEE Trans Intelligent Transp Syst* 17(12):3434–3445. <https://doi.org/10.1109/TITS.2016.2552248>
14. Nyamawe AS, Mbooso EC (2014) Road safety: adoption of ICT for tracking vehicles' over-speeding in Tanzania. *Int J Comput Appl* 96(16): 12–15. <https://doi.org/10.5120/16877-6876>
15. Pawłowicz B, Salach M, Trybus B (2019) Infrastructure of RFID-based smart city traffic control system. In: *Conference on automation*. Springer, Cham, pp 186–198. [https://doi.org/10.1007/978-3-030-13273-6\\_19](https://doi.org/10.1007/978-3-030-13273-6_19)
16. Byshov N, Simdiankin A, Uspensky I (2017) Method of traffic safety enhancement with use of RFID technologies and its implementation. *Transp Res Proc* 20: 107–111. <https://doi.org/10.1016/j.trpro.2017.01.030>
17. United Nations (UN) (2020) The 2030 Agenda for Sustainable Development. Online available: [https://ec.europa.eu/environment/sustainable-development/SDGs/index\\_en.htm](https://ec.europa.eu/environment/sustainable-development/SDGs/index_en.htm). Accessed on 19 Dec 2020
18. David S, Bill E, Tim L (2021) Urban Mobility Report 2019. Available online: <https://mobility.tamu.edu/umr/report/>. Accessed on 21 Mar 2021
19. World Road Association (2020) Road safety manual: a manual for practitioners and decision-makers on implementing safe system infrastructure. Paris; 2015. Online Available: <https://roadsafety.piarc.org/en>. Accessed on 27 Dec 2020
20. Global Plan for the Decade of Action for Road Safety 2011–2020. Online available: [https://www.who.int/roadsafety/decade\\_of\\_action/plan/global\\_plan\\_decade.pdf](https://www.who.int/roadsafety/decade_of_action/plan/global_plan_decade.pdf). Accessed on 25 May 2020
21. Mihyeon Jeon C, Amekudzi A (2005) Addressing sustainability in transportation systems: definitions, indicators, and metrics. *J Infrastructure Syst* 11(1):31–50
22. Ministry of Infrastructure. Transport Sector Strategic Plan for the National Strategy for Transformation (NST1). Online available: <https://www.mininfra.gov.rw/index.php?id=transportstrategicplan>. Accessed on 05 May 2021
23. Government of Rwanda. Smart Rwanda Master Plan. Available online: [http://minict.gov.rw/fileadmin/Documents/Strategy/SMART\\_RWANDA\\_MASTER\\_PLAN\\_FINAL.pdf](http://minict.gov.rw/fileadmin/Documents/Strategy/SMART_RWANDA_MASTER_PLAN_FINAL.pdf) (accessed on 10 October 2020)
24. National ICT Strategy and Plan, NICI-2015. [http://www.rdb.rw/uploads/tx\\_sbdownloader/NICI\\_III.pdf](http://www.rdb.rw/uploads/tx_sbdownloader/NICI_III.pdf) (accessed on 2 February 2021)

# Part of Speech (PoS) Tagging for Konkani Language Using HMM



Annie Rajan  and Ambuja Salgaonkar

**Abstract** For the first time, a part of speech (PoS) tagger tool using hidden Markov model (HMM) for Konkani the official language of the Indian state of Goa that is also spoken along the nearby coastal region has been built. For this study, over 100,000 PoS tagged Konkani sentences were used for the first time. The f-scores for HMM are 76.75% for training data and 71.29% for test data. To place our research in perspective, we provide a list of references of PoS tagging for Indian languages specifically using HMM.

**Keywords** Part of speech tagging · Konkani · Hidden Markov model · Part of speech tagger · Computing methodologies · Machine learning

## 1 Introduction

The process of annotating a part of speech to each word in a sentence is called part of speech tagging (PoS tagging). It is also known as word-category-disambiguation or grammatical tagging. The rules in a rule-based tagger are initially handcrafted by manual analysis of the data [1–4], and the tagging decision is based on the likelihood of the given word having that tag in a given corpus [5]. The methods proposed for PoS tagging include hidden Markov model (HMM) [6], support vector machine (SVM) [7], decision tree [8], maximum entropy [9] conditional random field (CRF) [10] and deep learning [11, 12].

The research results presented in this paper are that the hidden Markov model (HMM) gave f-score 76.75% and accuracy 99.20% after training, when employed on a set of 108,546 sentences of a Konkani database developed by the Indian Languages Corpora Initiative (ILCI) [13].

---

A. Rajan (✉)

Dhempe College of Arts and Science, Panaji 403001, India

A. Salgaonkar

Department of Computer Science, University of Mumbai, Mumbai 400032, India

e-mail: [ambujas@udcs.mu.ac.in](mailto:ambujas@udcs.mu.ac.in)

## 2 Literature Survey

For an appraisal of our PoS tagging experiments in Konkani, comparable work in various Indian languages is summarized: Table 1 describes data sets, and a performance result when HMM is employed.

Research initiatives in Konkani include the following: An HMM based PoS tagger using the NLTK tool kit employs a corpus of 50,000 Konkani sentences collected by ILCI. The reported accuracy is 73.82% [27]. A PoS tagger has been built by employing 250,000 sentences from the healthcare domain and 20,000 sentences from the tourism domain, collected from ILCI. 5000 sentences from health and

**Table 1** PoS tagging in Indian languages using HMM

S. No.	Language	Size of data set (words)	Performance measure accuracy (%)
1	Bengali [14]	100,000	89.00
2	Assamese [15]	271,890	89.21
3	Sinhala [16]	90,551	90.00
4	Punjabi [17]	26,479	90.11
5	Malayalam [18]	15,245	90.50
6	Telugu [19]	211,357	91.23
7	Manipuri [20]	2000	92.00
8	Gujarati [21]	351	92.87
9	Hindi [22]	15,562	93.45
10	Marathi [23]	25,744	93.82
11	Kannada [24]	80,000	94.62
12	Tamil [25]	43,678	95.92
13	Nepali [26]	150,839	96.00

**Table 2** Ten-fold validation output of training data for HMM

S. No	Cross validation	Recall (%)	Precision (%)	F-score (%)	Accuracy (%)
1	Run 1	78.51	75.42	76.70	99.20
2	Run 2	78.45	75.36	76.65	99.19
3	Run 3	78.56	75.37	76.68	99.20
4	Run 4	78.79	75.70	76.98	99.20
5	Run 5	78.81	75.51	76.88	99.19
6	Run 6	78.62	75.32	76.69	99.20
7	Run 7	78.56	75.34	76.68	99.20
8	Run 8	78.30	75.27	76.50	99.19
9	Run 9	78.45	75.29	76.60	99.20
10	Run 10	78.59	75.51	76.77	99.19

**Table 3** Five-fold validation output of test data for HMM

S. No.	Cross validation	Recall (%)	Precision (%)	F-score (%)	Accuracy (%)
1	Run 1	72.91	71.41	71.83	98.87
2	Run 2	73.18	71.43	71.94	98.87
3	Run 3	73.31	71.56	72.09	98.87
4	Run 4	72.95	71.41	71.84	98.86
5	Run 5	71.74	70.60	70.84	98.88

1000 sentences from tourism domains were tested. The tagger demonstrated accuracy of 89.00% and 87.00%, respectively [28]. An automatic POS tagger is trained by employing a standard data set of 5088 code-mixed Konkani-English sentences and tested with about 4000 Facebook and Twitter posts. A support vector machine (SVM) based classifier and conditional random field (CRF) based classifier have demonstrated an accuracy of 91.53% and 90.59%, respectively, for Facebook posts, and 92.53% and 91.47% for Twitter posts. That is, SVM slightly outperformed CRF [29]. Architecture is presented for a PoS tagger for Konkani that employs HMM in its basis and recommends the Viterbi algorithm to find the best tag for a given word [30].

### 3 Brief History of Konkani Language and its Speakers

Konkan is the name given to the central portion of the narrow coastal strip between the Arabian Sea and Western Ghats in peninsular India. The region is about 800 km long and about 45–75 km wide. The Konkani language is spoken along a total of about 600 km of the coast, along the central and southern parts of Konkan, and further south. As such, the geographical extent of the main Konkani speaking areas is relatively small. Moreover, it is not the only language in the region; the major Indian languages of Marathi and Kannada are also spoken alongside. It is the language of a population of 2.2 million, spread across neighbouring states of Goa, Karnataka, Maharashtra and Kerala.

There is no definite standard for the written or spoken Konkani language. Users and lovers of Konkani wrote in their own language or dialect, and in the local script. In the year 1975, the Sahitya Akademi (the national Academy of letters) declared Konkani as an independent language, rather than a dialect of Marathi. Today, it is one of the 22 languages officially recognized in the Indian constitution with Devanagari is its script. Organizations like Konkani Bhasha Mandal, Konkani Sahitya Parishad, Goa Konkani Akademi and Thomas Stephens Konkani Kendra are working towards promoting this language.

**Table 4** Word count of misclassified words using HMM

S. No.	Tag	Incorrectly tagged	Incorrectly tagged (%)
1	N_NN	22,595	19.14
2	JJ	13,931	11.80
3	V_VM_VNF	11,006	9.32
4	V_VM_VF	10,278	8.71
5	RB	9774	8.28
6	DM_DMD	6671	5.65
7	N_NST	5382	4.56
8	PR_PRP	4963	4.20
9	N_NNP	4887	4.14
10	PSP	3709	3.14
11	CC_CCS	2788	2.36
12	V_VAUX_VF	2783	2.33
13	DM_DMR	2755	2.06
14	QT_QTF	2428	1.36
15	PR_PRL	1607	1.17
16	RP_INTF	1385	1.17
17	RD_UNK	1384	0.87
18	QT_QTC	1026	0.76
19	PR_PRQ	893	0.76
20	CC_CCS_UT	893	0.76
21	CC_CCD	867	0.73
22	V_VM_VINF	857	0.73
23	V_VM_VNG	769	0.65
24	DM_DMQ	685	0.58
25	DM_DMI	640	0.54
26	RP_NEG	573	0.49
27	PR_PRI	487	0.41
28	RP_RPD	485	0.41
29	QT_QTO	461	0.39
30	PR_PRF	359	0.30
31	V_VAUX_VNF	225	0.19
32	V_VNG	222	0.19
33	PR_PRC	94	0.08
34	RP_INJ	47	0.07
35	V_VM	40	0.03
36	RD_ECH	38	0.03

(continued)

**Table 4** (continued)

S. No.	Tag	Incorrectly tagged	Incorrectly tagged (%)
37	RD_SYM	32	0.03
38	RP_CL	30	0.03
39	RD_RDF	18	0.02

**Table 5** Examples of words tagged incorrectly by HMM

Sr No.	Word	Tagged in gold-en standard	Incorrectly tagged by HMM
1	līp (लीप)	JJ	RB
2	əʃi (अशे)	RB	CC_CCS_UT
3	g <sup>he</sup> vʃe (घेवचे)	V_VM_VNF	V_VAUX_VF
4	kəʃe (करचे)	DM_DMD	V_VM_VF
5	kufik (कुश क)	PSP	RB

## 4 Data Set

A manually annotated corpus from ILCI in the domains of tourism, health, entertainment, agriculture and history with a total of 108,546 sentences is used for the experiments. Initially, the PoS tagger is trained on 86,837 sentences (1,006,177), and the test data has 21,709 sentences (125,389 words). The number of tags for the two data set is 39, whilst the BIS tag set for Konkani has only 29 tags [31].

## 5 Experiments

In an HMM, the exact sequence of states that a process  $X$  generates is unknown or hidden. As the process transitions to a next state, an output symbol is emitted; this is the observable process  $Y$ . Exactly which next state is entered, and which output is emitted, is governed by probability distributions. In HMM for a set of states,  $S = \{s_1, s_2, s_3, \dots, s_n\}$  and a set of outputs  $O = \{o_1, o_2, o_3, \dots, o_m\}$ , the model  $M = (A, B, \pi)$  is characterized by the transition probability matrix  $A = P(t_i | w_i)$ , emission probability or observed probability matrix  $B = P(o_k | S_i)$ , where  $V_k$  is the hidden state and  $\pi = P(S_i)$  is the vector of probabilities of initial states. Training the HMM, for a given observation sequence and corresponding state sequence, means estimating the transition matrix  $A$ , emission matrix  $B$  and initial probabilities  $\pi$  of the HMM. The following equations [32] explain the same

$$P(t_i | w_i) = P(t_i | t_{i-1}) \cdot P(t_{i+1} | t_i) \cdot P(w_i | t_i) \quad (1)$$





## 7 Discussion

The number of misclassified words for the HMM algorithm is 118,067 or 11.7% of the training data set of 100,6088. Table 4 contains the list of tags and the misclassified word count.

Observe from Table 4 that tags N\_NN, JJ, V\_VM\_VF, RB have high errors. The first row of Table 4 indicates that 22,595 words should have been tagged as nouns, but have not been tagged correctly. Words tagged incorrectly by HMM are shown in Table 5.

## 8 Conclusion

In this paper, we have presented POS tagger using HMM for Konkani. Whilst HMM has been employed earlier for this purpose [27–30] this is the first time 108,546 POS tagged sentences has been used and a POS tagger tool has been developed for Konkani. F-scores of 71.29% and using HMM and the error rates were 11.7%. N\_NN, RB, V\_VM\_VF, V\_VM\_VNF are the tags which resulted in high error rates. A detailed analysis of this situation will help in improving the f-score of the PoS tagger. Combining morphological analysis with PoS tagging is likely to help in making corrections for the misclassified tags. Future studies in morphological analysis and PoS tagging can be attempted for this purpose.

## References

1. Brill E (1992) A simple rule-based part of speech tagger. In: Proceedings of the third conference on applied natural language processing, Stroudsburg, PA, USA, pp 152–155
2. Brill E (1993) Automatic grammar induction and parsing free text: a transformation-based approach. In: Proceedings of the workshop on human language technology, Stroudsburg, PA, USA, pp 237–242
3. Brill E (1994) Some advances in transformation-based part of speech tagging. In: Proceedings of the twelfth national conference on artificial intelligence, Menlo Park, CA, USA, pp 722–727
4. Brill E (1995) Transformation-based error-driven learning and natural language processing: a case study in part-of- speech tagging. *Comput Linguist* 21(4):543–565
5. Doug C, Kupiec J, Pedersen J, Sibun P (1992) A practical part-of-speech tagger. In: Proceedings of the third conference on applied natural language processing, Stroudsburg, PA, USA, pp 133–140
6. Brants T (2000) TnT—a statistical part-of-speech tagger. In: Proceedings of ANLP-2000, Seattle, WA, arXiv preprint cs/0003055
7. Cortes C, Vapnik V (1995) Support-Vector networks. *Machine Learning, Springer* 20(3):273–297
8. Safavian SR, Landgrebe D (1991) A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern* 21(3):660–674
9. Berger AL, Della Pietra, Della Pietra VJ (1996) A maximum entropy approach to natural language processing. *Comput Linguist ACM* 22(1):39–71

10. Lafferty JD, McCallum A, Pereira FCN (2001) Conditional random fields—probabilistic models for segmenting and labeling sequence data. In: Proceedings of the eighteenth international conference on machine learning, San Francisco, CA, USA, pp 282–289
11. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Computation*, Massachusetts Institute of Technology (MIT), US 9(8):1735–1780
12. Jordan MI (1990) Attractor dynamics and parallelism in a connectionist sequential machine. *Artif Neural Netw Concept Learn IEEE* 112–127
13. Technology Development for Indian Language Programme, Ministry of Electronics and Information Technology (MeitY), Government of India, <http://tdil.meity.gov.in/>
14. Dandapat S, Sarkar S (2006) Part of speech tagging for bengali with hidden markov model. In: Proceeding of the NLP AI machine learning competition, Mumbai, India
15. Daimary SK, Goyal V, Barbora M, Singh U (2018) Development of part of speech tagger for assamese using HMM. *Int J Synth Emotions* 9(1):23–32
16. AJPMP Jayaweera, Dias NGJ (2014) Hidden markov model based part of speech tagger for Sinhala language, arXiv preprint [arXiv:1407.2989](https://arxiv.org/abs/1407.2989)
17. Sharma SK, Lehal GS (2011) Using hidden markov model to improve the accuracy of Punjabi PoS tagger. In: IEEE International conference on computer science and automation engineering, Shanghai, China, vol. 2, pp 697–701
18. Jayan JP, Rajeev RR (2011) Parts of speech tagger and Chunker for malayalam—statistical approach. *Comput Eng Intell Syst* (3)2
19. Gadde P, Yeleti MV (2008) Improving statistical PoS tagging using linguistic feature for Hindi and Telugu. In: Proceedings of the international conference on natural language processing, Hyderabad, India
20. Singha KR, Purkayastha BS, Singha KD (2012) Part of speech tagging in Manipuri with hidden markov model. *Int J Comput Sci Issues* 9(6):146–149
21. Yajnik A, Prajapati M (2017) Part of speech tagging using statistical approach for gujarati text. *Int J Appl Res Sci Eng*
22. Singh S, Gupta K, Srivastava M, Bhattacharya P (2006) Morphological richness offsets resource demand—experiences in constructing a POS tagger for hindi. In: Proceedings of the COLING main conference poster session, Sydney, Australia, pp 779–786
23. Singh J, Joshi N, Mathur I (2006) Development of marathi part of speech tagger using statistical approach. In: Proceeding of international conference on advances in computing, communications and informatics, IEEE, Mysore, India, pp 1554–1559
24. Chandramma D, Pareek PK (2018) Fast and accurate parts of speech tagging for kannada-telugu pair. *Int J Appl Eng Res* 13(10):7857–7867
25. Pandian S, Geetha TV (2008) Morpheme based language model for tamil part of speech tagging. *Polibits, México* 38:19–25
26. Paul A, Purkayastha BS, Sarkar S (2015) Hidden markov model based part of speech tagging for nepali language. In: International symposium on advanced computing and communication, Silchar, India, pp 149–156
27. Rajan A (2016) Design and implementation of a PoS tagger for Konkani using NLTK, MPhil Thesis, Goa University, India (unpublished) (2016)
28. Kane MMP (2017) Part of speech tagging for Konkani corpus. *Int J Eng Res Comput Sci Eng* 4(6):203–208
29. Phadte A, Thakkar G (2017) Towards normalising Konkani-English code-mixed social media text. In: Proceedings of the 14th international conference on natural language processing, Kolkata, India, pp 85–94
30. Khorjuvenkar DNP, Ainapurkar M, Chagas S (2018) Part of speech tagging for konkani language. In: Second international conference on computing methodologies and communication, IEEE, Erode, India
31. Sardesai M, Pawar J, Vaz E, Walawalikar S (2012) BIS Annotation standards with reference to Konkani language. In: Proceedings of the 3rd workshop on south and southeast asian natural language processing, Mumbai, India, pp 145–152

32. Joshi N, Darbari H, Mathur I (2013) HMM based PoS tagger for hindi. In: Third international conference on computer science & information technology, Bangalore, India, pp 341–349
33. Rajan A (2021) Part of speech (PoS) tagging for Konkani language using HMM

# Construction of Sports Articles Using Audio Commentary



**Nilkamal More, Priya Fernandes, Nachiket Bhuta, Riya Suri,  
and Akshita Patil**

**Abstract** In the current times, a plethora of the knowledge one has makes one realize the importance of summarized information which makes it easier to refer the same amount of knowledge in a shorter and concise manner. This paper focuses on generating sports articles for newspapers, magazines, and the likes of audio commentary to text conversion aspect as well of the matches associating cricket tournaments. By considering the varied format of data as an input, the project is able to deliver a matter outline of the commentary within the variability of an article as output. The project captures the idea of generation of summarized articles. It also revolves around the capturing of the idea of the speech which is nothing more than a collection of a variety of the sound waves generated by our vocal chords as air vibrates around them to the text conversion to capture an essence of the same. A microphone records the sound waves, which are then converted into an electrical signal. The model's accuracy will improve over time as it is trained. Speech recognition is accomplished by acoustic and language processing algorithms. Another aspect of the project is the development of summarized text, which is a way of condensing a long piece of text into a shorter version that we can use in our deep learning model. Reducing the initial text's size while retaining key informational elements and, as a result, the content's meaning and essence is preserved.

**Keywords** Speech-to-text · Text summarization · Seq2Seq architecture · TensorFlow · Sports · Commentaries · Articles · Audio

---

N. More (✉) · P. Fernandes · N. Bhuta · R. Suri · A. Patil

Department of Information Technology, K. J. Somaiya College of Engineering, Mumbai, India  
e-mail: [neelkamalsurve@somaiya.edu](mailto:neelkamalsurve@somaiya.edu)

P. Fernandes

e-mail: [priya.f@somaiya.edu](mailto:priya.f@somaiya.edu)

N. Bhuta

e-mail: [nachiket.bhuta@somaiya.edu](mailto:nachiket.bhuta@somaiya.edu)

R. Suri

e-mail: [riya.suri@somaiya.edu](mailto:riya.suri@somaiya.edu)

A. Patil

e-mail: [akshita.patil@somaiya.edu](mailto:akshita.patil@somaiya.edu)

# 1 Introduction

## 1.1 Problem Definition

The idea of this paper focuses on generating sports articles for newspapers, magazines, and the like via varied data for audio commentary of the matches and tournaments for which we intent to propose a system for a quick overview and analysis of the different kinds of sports articles present by summarizing the data. By taking audio as input, the project will be able to deliver a textual summary of the commentary in the form of an article as output. We will be implementing the same by making our own deep learning model. First, we will be converting the speech-to-text files. Audio is a variety of the sound waves generated by our vocal chords as air vibrates around them to the text conversion to capture an essence of the same. A microphone records the sound waves, which are then converted into an electrical signal. The model's accuracy will improve over time as it is trained. Speech recognition is accomplished by acoustic and language processing algorithms. Another aspect of the project is the development of summarized text, which is a way of condensing a long piece of text into a shorter version that we can use in our deep learning model. Reducing the initial text's size while retaining key informational elements and, as a result, the content's meaning and essence is preserved.

## 1.2 Scope

Any human-machine connectivity ecosystem also includes speech and natural language processing. The project highlights recent advances in artificial intelligence, factual information mining, and example identification, all of which would help to make discourse interfaces more versatile and inevitable. The changing specifications for discourse interfaces often warn of possible roadblocks to acoustically characteristic interfaces' efficient implementation. Finally, specialized advances and research projects for elite ongoing appreciation that will fully change how people communicate with their processing gadgets.

## 1.3 Salient Contribution

We will start from the ground up and build our own deep learning model. The project will be able to show the immense potential of summarized knowledge and how it helps many individuals to get valuable information that is fundamental to them by discussing the different aspects of the project that are important to them. This would have a clear view of the outlook of distilled papers as a means to represent important facts while maintaining the meaning of the knowledge communicated to

the client. The project's strength stems from the fact that manual summarization takes a long time, so designing a format for it will make for a better representation and more concise approach to feed the model information during the training phase and produce brief and to-the-point posts for the end user to read. The rouge metric score was increased from 0.35 to 0.41. We have also created our own dataset.

## 2 Literature Survey

In [1], the possibility of automatically collecting sports news from live text commentary scripts was investigated. This task is classified as a type of text summarization that relies on sentence extraction as a preliminary study. The challenge is designed using a supervised learning to rate framework that incorporates both standard sentence features for typical text summarization and novel task-specific features. To solve the problem of local redundancy, we propose a probabilistic sentence selection algorithm. Experiments based on data from live football commentary scripts and associated sports coverage have shown that this obstacle can be overcome. In a variety of ways, our methods outperform other baseline methods, showing that they are indeed appropriate for this purpose.

In [2], authors want to check whether it is possible to generate news based on the broadcast a set of comments that describe the game in real-time and considered as a summarization problem, using extractive and abstract approaches and encoding and text augmentation techniques.

In [3], here are three main things to consider. First, we have shown that cricket text commentary is a credible data source that has historically gone unnoticed. Second, we have seen how this information can be used to consider the strengths and shortcomings of players. Third, using the text commentary, data necessitates a considerable amount of processing and representation work. This work can be seen in many ways. To learn the rules, we must weigh a number of factors, including the game's difficulty level, the condition of the pitch, and the match's location. These laws can be used to construct probabilistic machine learning models that can predict the outcome of any ball. Coaches and managers in cricket will use our research to help them plan action.

In [4], the CNN daily mail data collection was used to train the model, which was done with TensorFlow. The model earned a Rouge-1 score of 0.38, which is considered to be cutting edge. The model, according to our observations, excels at generating summaries from news stories, which is the data it is trained on. When given a text that is not breaking news, though, it also produces useful summaries, though they are more extractive.

In [5], data summarization is a natural language processing task that aims to shorten the length of a text document while preserving as much meaning as possible. This pedagogical essay looks at a number of recent deep learning architectures that have helped to advance the technology. In brief, we will discuss pointer networks, hierarchical transformers, and reinforcement learning. We conclude that a basic

understanding of Seq2Seq architecture and transformer networks is important in NLP.

### 3 Usage of Dataset

We have created our own dataset for the training and testing purpose of the sequence-to-sequence model. We have used Web scraping using BeautifulSoup4 and requests modules of Python to download dataset from <https://cricbuzz.com> and Cricbuzz API to fetch the match-related data such as ball-by-ball commentary of each and every match in an IPL, and we have downloaded the same for IPL seasons (2008–2020). Also, we have used <https://www.espncricinfo.com> to fetch the news and summary of a cricket match for the sequence-to-sequence model for training purpose.

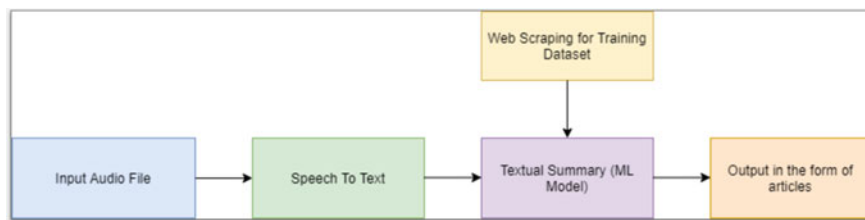
## 4 Implementation

See Fig. 1.

### 4.1 Problem

Two major issues arise from the Seq2Seq model with a focus on summarization:

1. In the abstracts, the details are frequently misrepresented (e.g., MI beat CSK by 131–129). This particularly applies to unusual or unknown words such as 131–129.
2. Resumes can often be recurring (e.g., Mumbai Indians beat Mumbai Indians beat Mumbai Indians beat Indians beat...).



**Fig. 1** Block diagram



## 4.2 *Problem-1 Explanation*

The sequence-to-sequence paradigm cannot handle copying a word  $w$  from the source text because it is too complicated. After the knowledge has been transferred through several layers of processing, the network must somehow restore the original concept (including mapping  $w$  to its word embedding).  $W$  is separate and unrelated in the network because  $w$  is an unfamiliar term in preparation and therefore has a weak word embedding (i.e., clustered with completely unrelated words).

The term is difficult to reproduce throughout the network; however,  $w$  is appropriate. RNN syntax systems, for example, often substitute a name other than Kohli—> Dhoni or a city of a different city (e.g., Delhi—> Mumbai). Since the term embedding seems to crowd in elements including Indian and feminine names, making reconstruction of the original word is difficult. This appears to be an overly complicated way to copy a simple method that is a concise description.

## 4.3 *Problem-2 Explanation*

An overreliance on decoder feedback from decoder data may be the explanation for this to be repeated in the context of the decoder's overreliance instead of storing long-term decoder data. This is seen by how often an unfinished repeat cycle starts with a single letter. For instance, a MI beat MI one-off error leads to a devastating MI beat MI beat MI beat.

## 4.4 *Solution to Problem-1*

Marker networks established copying facilitate: The point generation network is our solution to problem 1. You may also copy terms by marking from a source by generating words from the vocabulary of this hybrid network. We measure a concentration and vocabulary distribution as before. However, the probability  $p$ -genes generated between 0 and 1 are also known. This is the opportunity to generate a term instead of copying it from the source. The vocabulary  $M$  and distribution of treatment  $a$  have been weighed and merged to establish the final distribution.

The probability of creating the term  $w$  is actually the likelihood of generating it from the vocabulary (multiplied by the chance of generation) as well as the likelihood of pointing it somewhere in the source code (multiplied by the chance of pointing) (multiplied by the copying probability). The network created by the marker benefits from the Seq2Seq system:

1. The network of pointer generators allows easy copying of sentences from the source text. The network essentially has to concentrate on the term and increase the PGen scale.

- 2. The model of the pointer generator can also copy words from the original text not in the dictionary. This is a huge advantage because it helps one to deal with inflammatory words without having to restrict our vocabulary (which requires less computation and storage space).
- 3. The marker model created needs less training processes to achieve the same efficiency that the Seq2Seq method.

4.5 Solution to Problem-2

Using repeat reduction cover: We use a method called problem solving coverage 2 (repetitive summaries). The aim is to monitor the spread of attention already covered and penalize the network for returning to the same bits. The c cover vector is the sum of all t' focus distributions at every time of the decoder. In other words, it is equal to the amount of coverage a source name has received so far.

Finally, to penalize the overlap between the vector cover ct and the real focus distribution t, another fault word is added. This stops the network from debating topics that have already been discussed (Fig. 2).

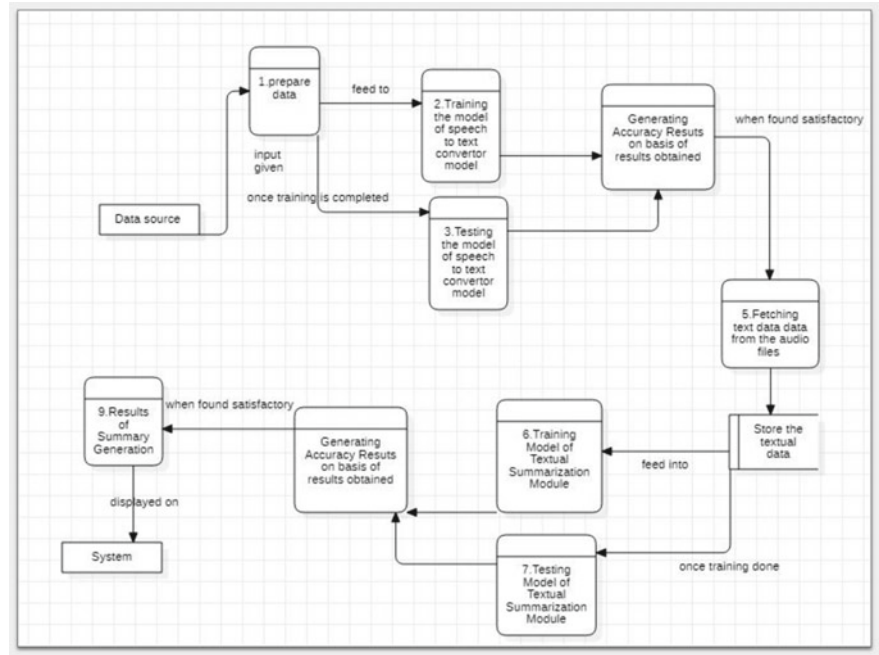


Fig. 2 DFD—level 1

## 5 Future Scope

While our answer abstract syntax is often terribly near to the initial language, it remains abstract. Higher abstraction, like a stronger compressive paraphrasing, is additionally a priority. Instead of opting to retrieve a vital secondary archive, the network continually neglects the core text of its supply. The network additionally assembles elements of the supply text, as Argentina claims to derail from Deutschland 2–0 though the contrary is true. Resumes in such phrases may work with an entire definition of function word, as an example, wherever a function word is employed to talk to a body while not a presentation. Interpretability, in my opinion, is the most important medium for prospective research. The methodology enlightens the recorder of neural network by exposing what the network “sees for,” allowing problems like replica and replica to be corrected.

## 6 Results

This prevents the network from discussing anything already addressed (thus summarizing). The simple sequence-to-sequence scheme states wrongly that the player is Indian—perhaps due to the Indian bias of the training results. The description provided by the simple sign continues to repeat itself, and we can see that the text of the source text also looks at the same sections. In the other side, there is no replication of the generator marker+cover model. If the network gives you a time to shorten or skip to another section of the document, PGen is big. Furthermore, the rouge was increased from 0.35 to 0.4135.

## 7 Conclusion

We have implemented this model using TensorFlow and trained it our own dataset. The model obtained a Rouge score of 0.41. The machine, according to our observations, excels at constructing summaries from news stories, which is the data it was trained on. When given a text that is not breaking news, it still produces effective summaries, but they are more extractive in character.

## References

1. Zhang J, Yao JG, Wan X (2016) Towards constructing sports news from live text commentary. In: Proceedings of the 54th annual meeting of the association for computational linguistics, pp 1361–1371

2. Porplenko D, Malykh V (2020) Generation of sport news articles from match text commentary. A master thesis from Ukrainian Catholic University
3. Behera SR, Agrawal P, Awekar A, Vijaya Saradhi V (2019) Mining strengths and weaknesses of cricket players using short text commentary. In: 2019 18th IEEE international conference on machine learning and applications (ICMLA)
4. Dwivedi P (2019) Text summarization using deep learning
5. Lemberger P (2020) Deep learning models for automatic summarization

# Driver Drowsiness Detection Using Eye Aspect Ratio (EAR), Mouth Aspect Ratio (MAR), and Driver Distraction Using Head Pose Estimation



T. V. N. S. R. Sri Mounika, P. H. Phanindra, N. V. V. N. Sai Charan,  
Y. Kranthi Kumar Reddy, and S. Govindu

**Abstract** Many people are getting injured and losing their lives because of road accidents. One of the main reasons among them is drowsy driving which is the main cause of road accidents and death. In major situations, fatigue is one of the key issues in road accidents. So, we should detect the fatigue in the initial stage, and this has become one of the trending research topics nowadays. Some of the important methods for detection of driver drowsiness are based on behavioral aspects of driver's face. By using the system, we can detect the face and determine the facial landmarks by which we can compute eye aspect ratio (EAR), mouth aspect ratio (MAR) to detect driver drowsiness based on adaptive threshold, and also using the head pose estimation, which checks the attention of the driver head with respect to the road, whether he is facing the road or not. When the system detects the driver having drowsiness, then it alerts the alarm. Head pose estimation majorly concentrates on three aspects, and they are pitch (used to find direction he is looking left or right), yaw (used to determine looking up or down), and roll (used to determine the rolling of face). In this way, we can detect drowsiness and distraction of driver.

**Keywords** Eye aspect ratio (EAR) · Mouth aspect ratio (MAR) · Driver drowsiness · Head movement · Yawning · Facial landmark · Monitoring system

## 1 Introduction

Road safety has become one of the key concerns to the people because of injuries and costing of lives of the people. There is no interest that is shown to reduce the accidents. There will be many reasons for the cause of accidents, 34% of accidents are occurred due to external environment, 24% of accidents are occurred due to attitude

---

T. V. N. S. R. Sri Mounika (✉) · P. H. Phanindra · N. V. V. N. Sai Charan ·  
Y. Kranthi Kumar Reddy

Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Krishna,  
Andhra Pradesh 521 230, India

S. Govindu

Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh 521 230, India

of driver, 20% of accidents are occurred due to driver drowsiness and fatigue, 17% of accidents occurred due to over speed of the vehicle, and remaining 5% are due to maintenance of vehicle. By observing the above statistics, driver fatigue is one of the important factors. From the year 1996 to 2006, number of vehicles has been increased by 20% which may also increase the number of accidents.

If a person is feeling drowsy, it impacts him to take decision in controlling the car. A person can be identified whether he is in drowsy state if he blinking his eyes frequently, yawning, misreading of traffic signs, etc. If a person is having less than 4 h of sleep, then he is 10 times more likely to involve in an accident. Sometimes the accidents may also occur due to distraction of the driver. We have observed that sometimes the driver will turn to his left side or right side. For watching anything outside the vehicle or talking to the person sitting beside him. To detect the driver's distraction, we will use a technique called as head pose estimation.

Head pose estimation is associated with visual gaze estimation that means the ability to determine the direction and focus of a person's eyes. Head pose estimation is mainly used in situations where we cannot determine the eyes of a person, and this occurs in low intensity or lowlight areas. By using both visual gaze estimation and head pose estimation, we can accurately predict the direction of the face and focus of eyes of the person.

## 2 Literature Survey

Driver drowsiness detection using percentage eye closure method (2020) used many varieties of research done on the drowsiness which are seen in driver like yawning, eye closure, and head movement. The proposed system focuses on eye closure by the effective method, named percentage eye closure (PERCLOS). The concept of this method is that it rejects the non-face easily and spends time on the face region.

Shailesh (S. Sangle, B. Rathore, R. Rathod, A. Yadav, and A. Yadav) in 2018 with the help of a camera fixed in dashboard that will forward the captured frames to the server of the Raspberry Pi configured in the vehicle. In this sever for detecting the face it will use Harr cascade classifier and for identifying landmarks of the face it will use dlib.

In order to detect drowsiness of drivers, numerous approaches are proposed. Here is an existing system proposed by Khalifa Al-Khalifa, Mohammed Shinoy, Rateb Jabbar, Mohammed Shinoy, Mohamed Kharbeche, Khalifa Al-Khalifa, Kamel Barkaoui, and Moez Krichen in 2018, detected real-time driver drowsiness using deep neural networks. They developed an Android application.

### 3 Proposed Work

Methodology of our project is as shown below. It mainly contains the five steps, and they are face detection, identifying facial landmarks, calculation of EAR and MAR, determining the distraction, and conclusion (Fig. 1).

We proposed a real-time monitoring system that will detect the drowsiness of the driver and distraction of the driver. Drowsiness of the driver can be detected by using eye aspect ratio (EAR) and mouth aspect ratio (MAR). The distraction is used to determine whether driver is looking right or looking left. For detecting the direction, we will be using a technique called head pose estimation.

#### 3.1 Input Data

For this system, we will use one camera in order to capture the frames of the driver. We will start capturing the frames when the driver starts driving the vehicle. Python provides various libraries for image and video processing. We will be using video stream method to capture the video and extract the frames from it.

#### 3.2 Face Detection

When we start this application, first it will start capturing the video. From the video, we will retrieve each single frame, and we will detect that the face is present in the particular frame. In order to detect the face in frame, we will be using one object detection approach called as Haar cascade classifier. It is one of the efficient approaches because it is having good ability in determining edge features and line features. After face detection, we will identify the facial landmarks.

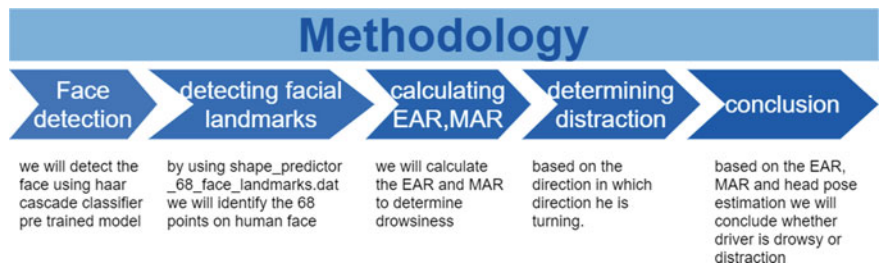
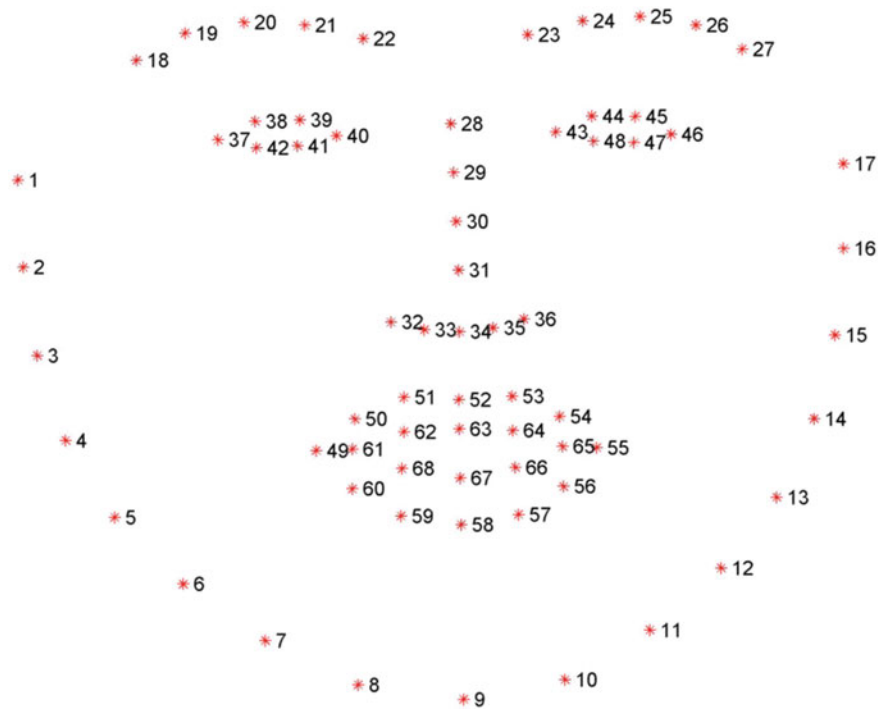


Fig. 1 Steps involved in methodology



**Fig. 2** Shows all 68 facial points

**3.3 Identifying Facial Landmarks**

To determine the facial landmarks, we will be using a library called as dlib. We will also be using a pre-trained model file called as shape\_predictor\_68\_face\_landmarks.dat which is used to detect all the 68 points on the face. The files with dat extensions are generic data file that stores the specific information relating to the program that created the file (Fig. 2).

**3.4 Eye Aspect Ratio (EAR) Calculation**

The **eye aspect ratio** is the estimation of the **eye** opening state. By using the above facial points, we know that points from 37 to 48 represent the eye coordinates. From 37–42, it represents the coordinates of right eye, and 43–48 represents the coordinates of left eye. The formula to calculate the EAR is as shown Figs. 3 and 4.

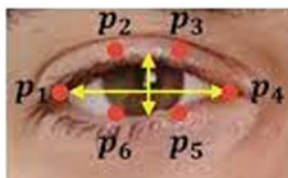
As shown in the above diagram, we can say that if the eyes are closed, the value of EAR will decrease. As shown in the figure, when eyes are closed, the value of  $||p2 - p6|| + ||p3 - p5||$  will decrease which results in decreasing the EAR value.



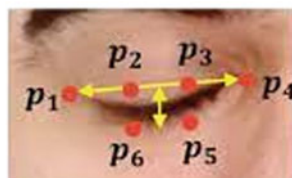
**Fig. 3** Eye aspect ratio formula

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

Eye aspect ratio will be larger and relatively constant over time when eye is open



Eye aspect ratio will be almost equal to zero when a blink occurs



**Fig. 4** Eye points

Initially, we will take a variable called as EAR\_THRESHOLD and assign some value to it. For every frame captured, we will calculate the eye aspect ratio, and if it is greater than the EAR\_THRESHOLD, then we will increment the counter. If the value of the counter reaches 40 means for 40 frames the value of EAR is greater than EAR\_THRESHOLD, now we can say that the person is drowsy.

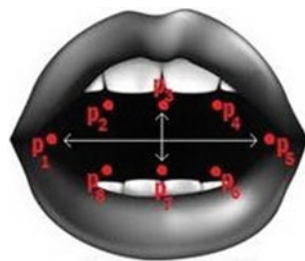
### 3.5 Mouth Aspect Ratio (MAR) Calculation

Mouth aspect ratio (MAR) is used to detect whether a person is yawning or not. By using the above facial points, we know that points from 49 to 68 represent the mouth coordinates. But we will use only eight points, and they are 61–68 points. The formula for mouth aspect ratio (MAR) calculation is as shown (Figs. 5 and 6).

**Fig. 5** Mouth aspect ratio formula

$$MAR = \frac{\|p_2 - p_8\| + \|p_3 - p_7\| + \|p_4 - p_6\|}{2\|p_1 - p_5\|}$$

**Fig. 6** Facial points of mouth



When we are feeling sleepy, we will automatically yawn. By knowing, if person is yawning or not, we can determine whether he is feeling drowsy or not? Sometimes driver will open his mouth while he is talking, but we will not consider it as yawning.

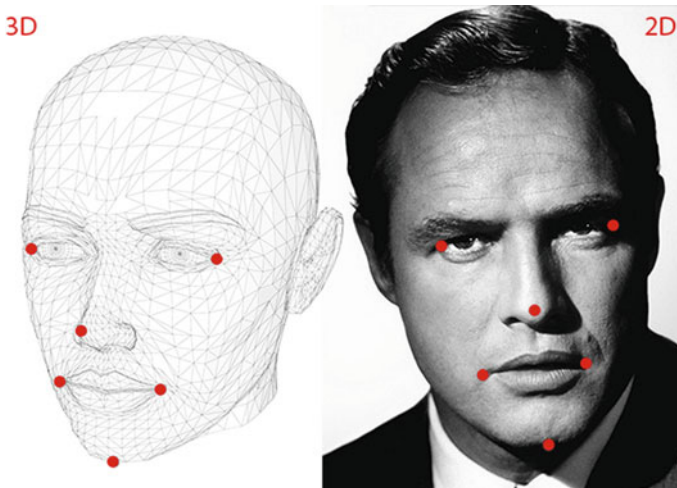
It is also similar to EAR. If the value of the calculated MAR is greater than the threshold value, we increase the counter variable. If the value of the counter variable also reaches to a particular value, we can say that the person is drowsy.

### 3.6 Head Pose Estimation

Head pose estimation is one of the challenging tasks in the field of computer vision. First, we need to locate all the facial landmarks. After finding the landmarks, we need to convert the facial points into 3D coordinates to find the inclinations.

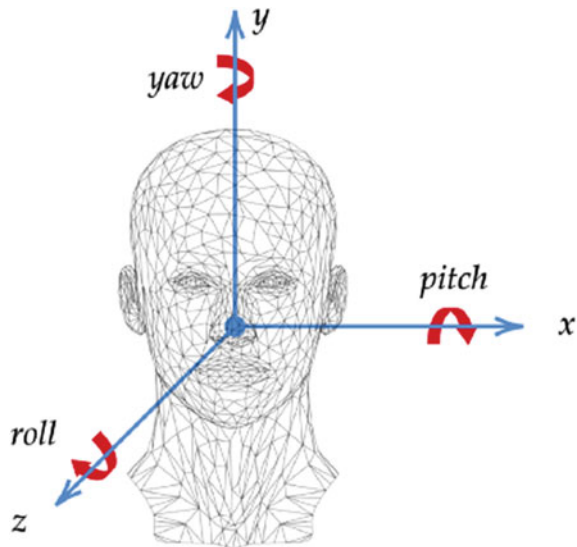
We need six points on the face to determine head pose, i.e., extreme left and right points of lip, chin, nose tip, right corner of right eye, and left corner of left eye. We should convert these points into 3D coordinates and find the translational vectors of nose tip (Figs. 7 and 8).

As shown in figure, we will get three directions in which human face can turn. Pitch(x) indicates the angles when we are moving our face in up and down direction. Yaw(y) indicates the angles when we are turning our face in left and right direction. Roll(z) indicates the angles when we are tilting our face. In this way, head pose estimation is used in estimation of the driver distraction (Fig. 9).

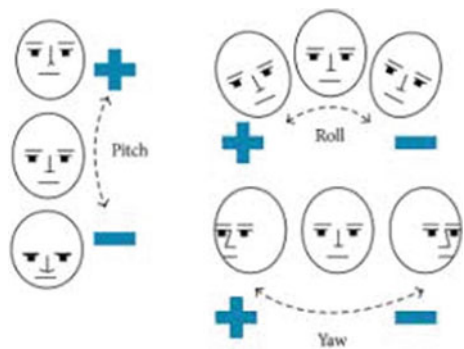


**Fig. 7** Shows six points required for head pose estimation

**Fig. 8** Three directions of human face



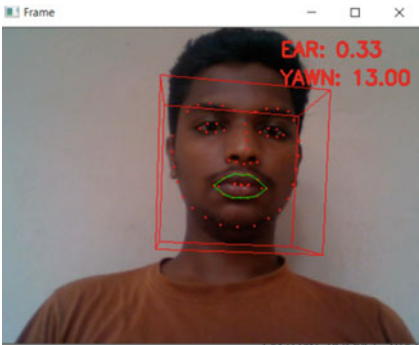
**Fig. 9** Yaw, pitch, and roll working



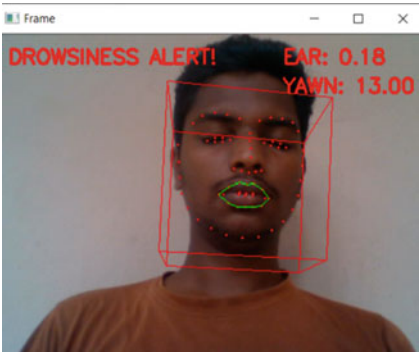
**4 Results**

By performing the above methodology, we have obtained the results as shown below. In the Fig. 9, it is showing the drivers face and displaying the eye aspect ratio and mouth aspect ratio. Both the EAR and MAR values are within the given thresholds, so it is not showing any drowsiness or distraction alerts. In Figs. 10 and 11, EAR and MAR are gone beyond the given thresholds, so it has shown drowsiness alert and yawn alert. In Fig. 12, the driver is distracting from focusing on the driving, so it is showing looking left side (Fig. 13).

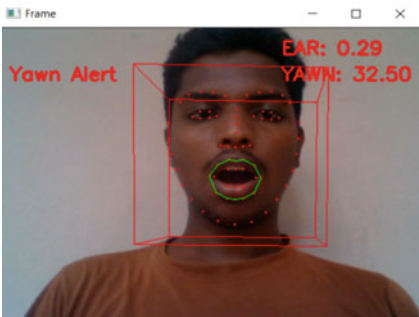
**Fig. 10** Calculating EAR and MAR



**Fig. 11** Detecting drowsiness



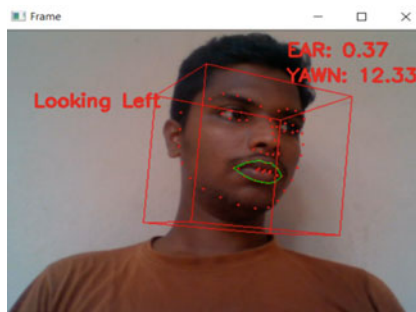
**Fig. 12** Detecting yawning



5 Conclusion

By observing the above results, we can conclude that by using this application, we can detect the drowsiness of the driver using eye aspect ratio, mouth aspect ratio, and driver distraction using head pose estimation. This application helps us in achieving less accident percentage that is occurring due to driver fatigue. Sometimes this application will not detect the face due to low intensity of light on the driver face.

**Fig. 13** Detecting distraction



We can overcome this problem by using night vision cameras which can detect the face even in the low intensity of light. When we identify the drowsiness of a person, we can send the information of the vehicle to nearby police station so that they will take necessary action.

## References

1. <https://www.nsc.org/road-safety/safety-topics/fatigued-driving>
2. <https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>
3. <https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/>
4. <https://www.pyimagesearch.com/2017/05/08/drowsiness-detection-opencv/>
5. <http://vision.fe.uni-lj.si/cvww2016/proceedings/papers/05.pdf>
6. <https://medium.com/analytics-vidhya/real-time-head-pose-estimation-with-opencv-and-dlib-e8dc10d62078>
7. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3356401](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3356401)
8. <https://www.instructables.com/Drowsiness-Alert-System/>
9. <https://bjrbe-journals.rtu.lv/article/view/bjrbe.2021-16.518>
10. [https://www.researchgate.net/publication/328911666\\_RealTime\\_Driver's\\_Drowsiness\\_Monitoring\\_Based\\_on\\_Dynamically\\_Varying\\_Threshold](https://www.researchgate.net/publication/328911666_RealTime_Driver's_Drowsiness_Monitoring_Based_on_Dynamically_Varying_Threshold)

# Frequency Reconfigurable Patch Antenna for 5G Applications



Sanjeev Kumar, Sneha Prasad, Shubham Goswami, Tsarina Patnaik, and Satyaki Tatte

**Abstract** With today's world growing wireless and with the emerging concept of 5th Generation (5G) technology, flexible antennas are being preferred over other existing technologies. The antenna design presented in this paper is a compact size multiband frequency reconfigurable antenna for wireless applications with simple line feed using FR-4 Epoxy as the substrate. The resonating frequency range for the antenna varies between 6 and 21 GHz for ON condition and between 14 and 26 GHz for the OFF condition. The shift in frequency is achieved using BAR63-02 V, pin diode, at different slots in the patch. The chosen material for the substrate is FR-4 Epoxy because of its versatile nature with a dielectric constant of 4.4 and thickness of 1.6 mm.

**Keywords** Frequency reconfigurable · Multiband · 5G · PIN diode · HFSS · Edge feed · Biasing circuit · Return loss · VSWR

## 1 Introduction

An antenna can be used for multiple purposes by simply making it frequency and pattern reconfigurable. Reconfigurable antenna is an antenna that has the capability of modifying features like frequency, radiation pattern, and polarization dynamically,

---

S. Kumar (✉) · S. Prasad · S. Goswami · T. Patnaik · S. Tatte  
Department of Electronics and Telecommunications, Symbiosis International (Deemed University), Pune, India  
e-mail: [sanjeevkumar@sitpune.edu.in](mailto:sanjeevkumar@sitpune.edu.in)

S. Prasad  
e-mail: [sneha.prasad@sitpune.edu.in](mailto:sneha.prasad@sitpune.edu.in)

S. Goswami  
e-mail: [shubham.goswami@sitpune.edu.in](mailto:shubham.goswami@sitpune.edu.in)

T. Patnaik  
e-mail: [tsarina.tsarina@sitpune.edu.in](mailto:tsarina.tsarina@sitpune.edu.in)

S. Tatte  
e-mail: [tatte.satyaki@sitpune.edu.in](mailto:tatte.satyaki@sitpune.edu.in)

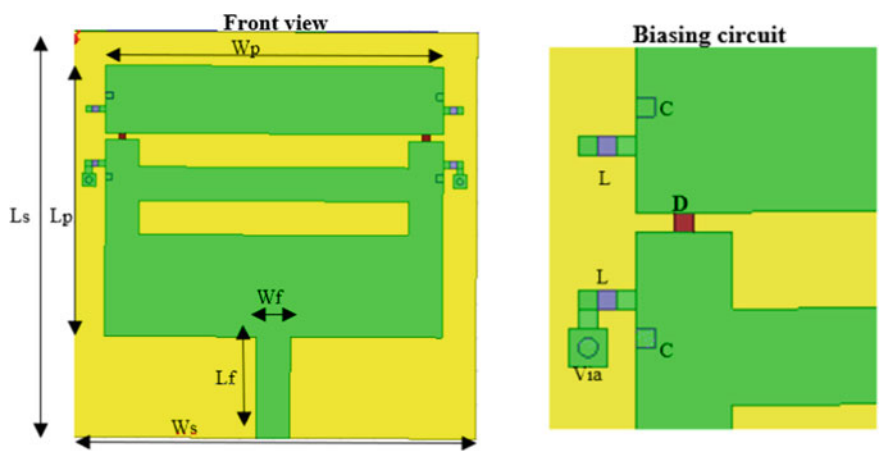
in a very controlled and reversible way. Various switching mechanisms can be used to achieve reconfigurability in an antenna by using PIN diodes, MEMS, or varactor diodes. Frequency reconfigurability may be obtained by smooth shift in between operating bands by changing the current length in the patch antenna. Reconfigurable designs have been developed by various researchers for multiple applications ranging from wireless systems to space technologies.

In [1], a coplanar waveguide-fed wideband antenna with reconfigurable rejection band characteristics is developed using MEMS switches. Liquid crystal polymer (LCP) was used to fabricate the antenna and operated in the frequency range 5 and 6 GHz. The antenna had a width of 26.88 mm and length of 20.7 mm. In [2], advantages and disadvantages of reconfigurable antenna are discussed. It further mentions in detail about various type of reconfigurable antenna and how PIN diode, varactor diode, MEMS configuration are used to achieve desired reconfigurability. In [3], a design for rectangular patch antenna with 5th generation application is proposed. The patch of the antenna is of  $3.0526 \text{ mm} \times 4.2352 \text{ mm}$  using Rogers RT5880 LZ with a height of 1.6. The antenna operates in the range of 26.5–28.8 GHz and was verified with the manufacturer's values. In [4], a novel unidirectional wideband antenna design with a shortened patch is presented. The proposed antenna works in the frequency range of 1.85–2.89 GHz while  $\tau$ -shaped strip feed line is used to excite the antenna. The antenna is of the dimension  $60 \text{ mm} \times 30 \text{ mm}$  and has more than 43.8% impedance bandwidth for  $\text{SWR} < 1.5$ . In [5], reconfigurable antenna using slots and PIN diode is proposed which switches polarization states. The resonating frequency for this antenna of 10 dB reflection coefficient is 2.4 MHz and has an axial ratio of 3 dB. Developments made in the patch antenna that shows frequency shift and has application in the L-Band in [6] showed better results. This antenna was developed using FR-4 substrate, inset feed and defected ground structure (DGS). The ground had concentric rings in the shape of a square. Frequency shift was achieved using PIN diodes and the antenna resonated at 1.61–1.82 MHz. In [7], a review on different kinds of antenna working in 5G has been described. Their performance, efficiency, and future scope are mentioned in a holistic approach. In [8], a frequency reconfigurable antenna of size  $30 \text{ mm} \times 25 \text{ mm} \times 2 \text{ mm}$  is proposed with 6 RF PIN diodes for reconfiguring frequency ranging from 4 to 8.5 GHz. The main highlight of this paper is that this reconfigurable antenna can resonate in single as well as multiband keeping the design unchanged. In [9], a simple rectangular patch antenna is designed to obtain both frequency and pattern reconfiguration. This antenna was operating at a frequency of 4.5–4.8 GHz/5.2–5.8 GHz with tilt of  $-30$  to  $+30^\circ$ . In [10], a prototype reconfigurable antenna was designed which operates in the S-band ranging from 3.4 to 3.8 GHz with peak return loss of  $-38 \text{ dBi}$ . The antenna was simulated using FR 4 as the substrate.

This paper presents an innovative design of a microstrip line fed patch antenna to operate in the 5G frequency band. Section 2 describes the antenna design and structure, and in Sect. 3, the simulated results have been discussed. The aim of the paper is to design and showcase the aspect of reconfigurability at 5G frequency with the use of a PIN diode and its biasing circuit.

## 2 Antenna Structure and Design

The proposed frequency reconfigurable antenna, as shown in Fig. 1, is designed using HFSS software. It is a rectangular patch separated into two sections. The first section has a slot of 36 mm x 2 mm, and the second section is separated by a small gap that is connected through two BAR63-02 V PIN diodes. The antenna patch dimension is 40 mm × 50 mm × 1.6 mm, and two PIN diodes are incorporated in the patch to achieve frequency shift since they act as switch for reconfiguration. It is an edge-fed rectangular patch antenna with a feedline length of 12 mm and feedline width of 3 mm. A low dielectric constant substrate is used for the design of the antenna. The substrate size of the proposed antenna is 60 mm × 60 mm, made up of FR-4 with thickness of 1.6 mm, dielectric constant 4.4, and loss tangent ranging between 0.02–0.03. The slot width of the antenna for incorporating the PIN diodes is 2 mm each. All the parameters used to implement this design of frequency reconfigurable antenna are mentioned in Table 1.



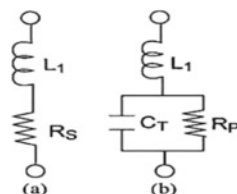
**Fig. 1** Front view and biasing circuit of the proposed frequency reconfigurable antenna

**Table 1** Parameters and dimensions of the proposed antenna

S. No.	Parameters	Dimension (mm)
1	Patch Width ( $W_p$ )	40
2	Patch Length ( $L_p$ )	50
3	Substrate Width ( $W_s$ )	60
4	Substrate Length ( $L_s$ )	60
5	Substrate Height ( $H_s$ )	1.6
6	Feed Width ( $W_f$ )	3
7	Feed Length ( $L_f$ )	12
8	Slot Width ( $S$ )	2



**Fig. 2** **a** ON state of pin diode, **b** OFF state of pin diode



The radiating patch of the proposed frequency reconfigurable antenna has 2 BAR63-02 V PIN diodes incorporated between the two sections of the patch as shown in Fig. 1. The antenna has edge feed with feedline width of 3 mm to achieve an impedance match of  $50\ \Omega$ . The design of the antenna along with biasing circuit that includes the blocking capacitor  $C = 20\text{pF}$ , RF choke inductor  $L = 33\text{nH}$  and via for ground is shown in Fig. 1. Blocking capacitor (C) works on the concept of blocking DC and thereby preventing spurious bias of the PIN diodes. The RF choke works as an inductor and is used here to allow DC current to flow through the circuit and to block AC current from passing through circuits. When diode is forward biased, any DC signal that changes the biasing of the diode is blocked by the capacitor C. While for RF choke inductor (L), any AC signal that changes the biasing of the diode is blocked.

Incorporating the PIN diodes between the two sections of the patch results in increasing and decreasing the overall length of the patch which has a direct impact on current length in the patch. This change in physical length of the patch alters the current length and consequently the resonant frequency. The equivalent circuit of the BAR63-02 V PIN diode is shown in Fig. 2. The frequency reconfigurability is achieved for this proposed antenna by switching ON/OFF, the two BAR63-02 V PIN diodes. This is incorporated in the slots by using lumped RLC open in HFSS. The datasheet of this diode mentions that in ON condition,  $L = 1\text{nH}$ ,  $R = 1.2\ \Omega$ , and in OFF condition,  $L = 1\text{nH}$ ,  $R = 20\text{kohm}$ , and  $C = 0.2\text{pF}$ . The blocking capacitor and RF choke inductor with values  $20\text{pF}$  and  $33\text{nH}$ , respectively, are used for biasing of these PIN diodes.

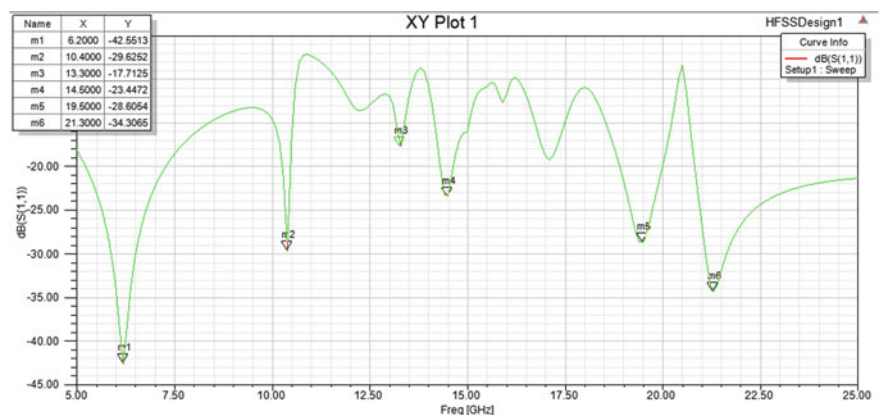
### 3 Simulated Results

The proposed frequency reconfigurable antenna is designed and simulated in HFSS software. The antenna can be reconfigured into different frequencies ranging from 6.2 GHz–21.3 GHz in ON condition and 14.1 GHz–26 GHz in OFF condition. Return loss, VSWR, radiation pattern, gain, electric field, and current distribution are measured in this software to realize the performance and efficiency of the frequency reconfigurable antenna. The s-parameter for both ON/OFF conditions is mentioned in Figs. 4 and 5, respectively. The shift in frequency is analyzed by keeping two PIN diodes, D1 and D2, in different modes of ON/OFF conditions. Table 2 shows different modes of the PIN diodes and the obtained resonating frequency in each case.

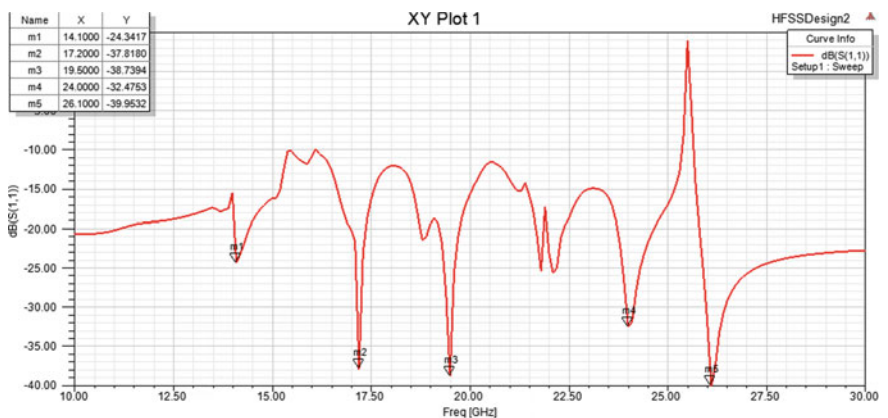
**Table 2** Modes of operation and resonating frequency

Modes	PIN diodes	Resonating frequency
1 ON condition	D1 = ON, D2 = ON	6.2,10.4,13.3,14.6,19.5,21.3
2 OFF condition	D1 = OFF, D2 = OFF	14.1,17.2,19.5,24,26

When both the diodes, D1 and D2, are in ON condition, the proposed antenna resonates at a frequency range of 6.2 GHz–21.3 GHz with a peak return loss of −42 dBi. However, when D1 and D2 are in OFF condition, the antenna operates at a frequency of 14.1 GHz–26 GHz with a peak return loss of −36.96 dBi. The s-parameter for both ON/OFF condition is mentioned in Figs. 3 and 4, respectively.



**Fig. 3** S-parameters for ON condition of the proposed frequency reconfigurable antenna



**Fig. 4** S-parameters for condition-OFF of the proposed frequency reconfigurable antenna

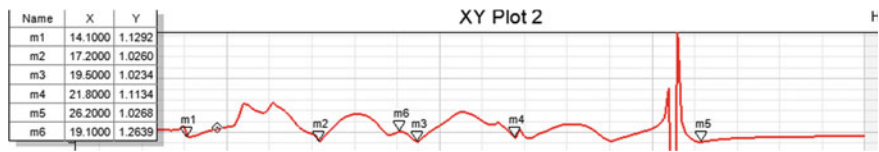


Fig. 5 VSWR for condition-OFF of the proposed frequency reconfigurable antenna

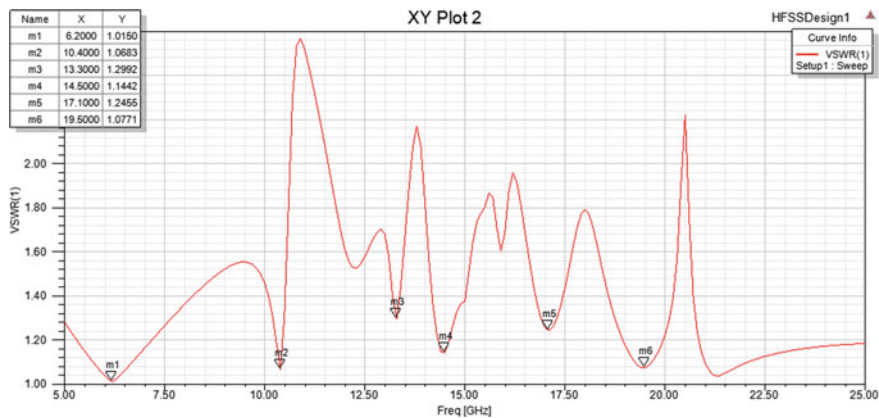


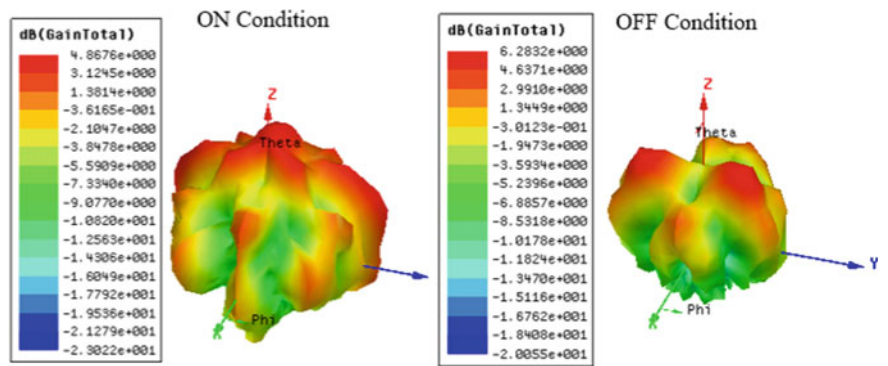
Fig. 6 VSWR for condition-ON of the proposed frequency reconfigurable antenna

The VSWR for the proposed frequency reconfigurable antenna can be viewed in Fig. 5 for OFF condition and Fig. 6 for ON condition. D1 and D2 in ON condition resonate at a frequency of 6.2, 10.4, 13.3, 14.6, 17.1, and 19.8 GHz and with VSWR of 1.01, 1.06, 1.29, 1.13, 1.24, 1.07 is showing Fig. 6 and in OFF condition resonates at frequency of 14.1, 17.2, 19.5, 21.8, 26.2, 19.1 GHz with VSWR of 1.12, 1.02, 1.02, 1.11, 1.02, 1.26 as shown in Fig. 5.

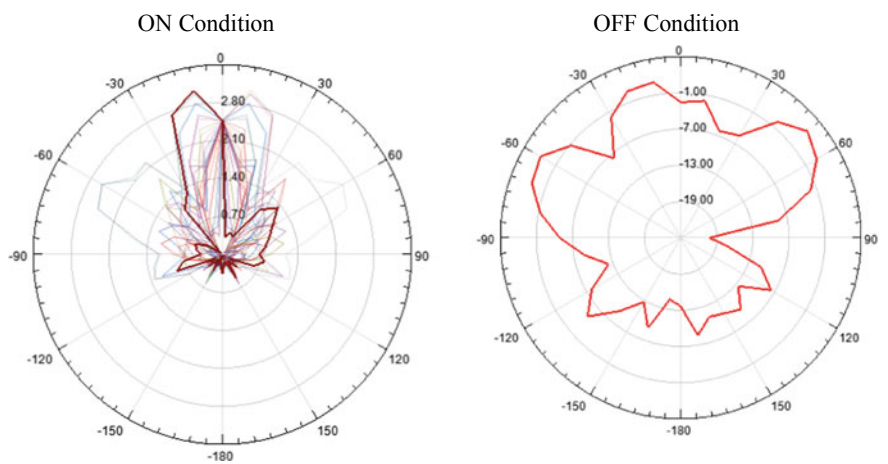
The gain measured for this frequency reconfigurable antenna can be viewed in Fig. 7 and is at a peak of 4.86 dBi when both D1 and D2 are in ON condition. However, the gain measured when D1 and D2 are in OFF condition range up to 6.28 dBi. These parameters can be viewed in Fig. 8. Slotting technique is used in the design to provide radiation pattern in a particular direction seen from Fig. 8, and most of the power is concentrated at z-direction.

For the proposed antenna, the radiation pattern is shown in Fig. 8, at  $\phi = 0$  degree for a frequency range of 20 GHz for ON condition and, the radiation pattern at  $\phi = 0$  degree for a frequency range of 15 GHz for OFF condition, respectively.

Performance of this frequency reconfigurable antenna by calculating the bandwidth, gain and efficiency percentage are observed. For OFF condition, the gain is 4.6 dBi, 4 dBi and efficiency is 85%, 82% for frequency of 6.2 GHz and 21.3 GHz, respectively. While for ON condition, the gain is 4.61 dBi, 3.8 dBi, and efficiency is 84%, 83% for frequency of 19.5 GHz and 26.1 GHz, respectively. The prototype of this antenna can be visualized as three different rectangular slots of different



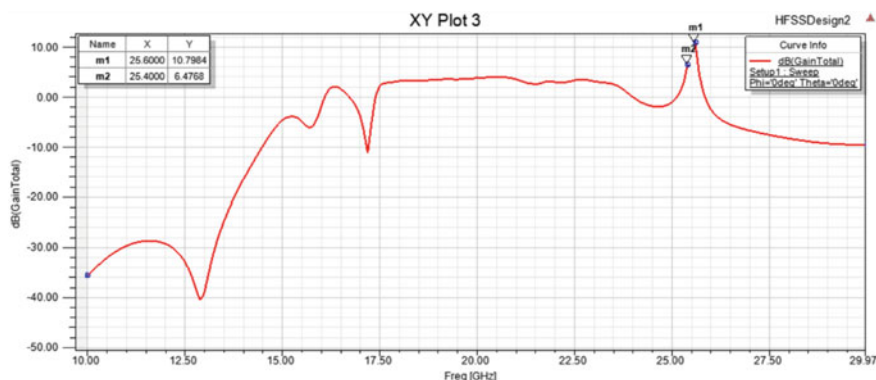
**Fig. 7** 3D Polar plot representing gain for condition-ON, OFF resp of the proposed frequency reconfigurable antenna



**Fig. 8** Radiation pattern for condition- ON, OFF resp of the proposed frequency reconfigurable antenna

lengths which are coupled through capacitors and can be linked with each other using diodes. Concept of electrical length is taken into the use so as to improve overall efficiency, and it's working capabilities were further enhanced by incorporating choking inductor and blocking capacitors.

The frequency versus gain plot for  $\phi = 0^\circ$  can be viewed in Fig. 9. This plot shows that we have two peaks at 25.6 GHz and 25.4 GHz which have a gain of 10.798 dBi and 6.478 dBi. This plot proves the 3D polar plot and efficiency that is obtained for this proposed antenna when the diodes are in forward bias, that is, ON condition.



**Fig. 9** Frequency versus gain for ON condition

## 4 Future Scope and Conclusion

The proposed design is a multiband, multi-application antenna which can cater several 5G frequencies in the range of (5–26 GHz) for telecommunication purpose, also can be used in remote sensing, advanced communication system such as satellite communication and aircraft navigation.

In this paper, a frequency reconfigurable patch antenna with edge feed has been discussed for 5G applications. The antenna resonates at a frequency of 6–21 GHz for ON condition and between 14 and 26 GHz for the OFF condition, and this shift is achieved by using pin diodes. The antenna provides a gain 4.8–6.8 dBi and has a good impedance bandwidth that is useful for future applications in the 5G arena.

## References

1. Symeon N, Nickolas K, George P, John P, Manos T (2009) UWB elliptical monopoles with a reconfigurable band notch using MEMS switches actuated without bias lines. *Antennas and propagation*. IEEE Trans 2242–2251
2. Haydar A, Salah M (2016) A study of reconfigurable multiband antenna for wireless application. *Int J New Technol Res* 2(5):2454–4116
3. Saeed R, Qunsheng C, Ishfaq H, Hisham K, Muhammad Z, Waseem N (2017) Design of rectangular patch antenna array for 5G wireless communication. In: *Progress in Electromagnetics Research Symposium*, pp. 1558–1562 (2017)
4. Lei G, Luk KM (2012) A Wideband Magneto-Electric Dipole Antenna. *IEEE Trans Antennas Propag* 60:4987–4991
5. Yang X, Shao B, Yang F, Elsherbeni AZ, Gong B (2012) A polarization reconfigurable patch antenna with loop slots on the ground plane. *IEEE Antennas Wirel Propag Lett* 69–72
6. Kumar S, Khandekar R, Waghmare PT (2019) Frequency reconfigurable patch antenna for L Band applications. *Int J Eng Adv Technol (IJEAT)* 9(1):2249–8958
7. Kumar S, Dixit A, Rajeshwari M, Raut H, Shevada L (2020) Fifth generation antennas: a comprehensive review of design and performance enhancement techniques. *IEEE Access*, pp 1–1

8. Chitra RJ, Velmurugan N (2014) Frequency reconfigurable antenna using PIN diodes. In: 20th National conference on communications, NCC 2014, pp 1–4
9. Yogeshwari S, Malathi K, Gulam M, Sangeetha V, Saffrine K, Sangeetha S et al (2017) A low profile frequency and pattern reconfigurable antenna. *IEEE Antennas Wirel Propag Lett* 1–1
10. Prakash V, Jothilakshmi, Shangaralavakusaraja S (2014) Design and implementation of reconfigurable microstrip patch Antenna for S-Band applications. *WSEAS Trans Commun* 178–185

# An Analysis of Future 5G Wireless Sensor Network



J. N. Swaminathan, M. Alagumeenaakshi, S. Umamaheswari, V. Jaikumar, and N. S. Kalyan Chakravathy

**Abstract** Wireless sensor networks (WSNs) are sensor nodes that are interconnected to communicate wirelessly to obtain data from the close environment. The upcoming technology development in the fields such as video surveillance, robot control, industrial automation, and so on can be achieved only with the help of WSNs. These distributed nodes will communicate wirelessly among them and receive the data from the user and transfer it to the main node. WSNs play a major role in wide range of applications including civilian applications. Though it has significant research interest, still, some of the problems such as memory limitation, security, battery issues, bandwidth, and computing should be rectified. In this review, the practical and theoretical achievements in the field of WSNs have been reviewed, and their applications towards the engineering field are also discussed. Along with that, the problems to overcome and the future perspectives on WSNs are also discussed.

**Keywords** Wireless sensor networks · Smart engineering · Sensors · Nodes

## 1 Introduction

Significant technology advancement made all the human beings to live their life sophisticated. They can monitor all the data using the wireless gadgets. In these applications, WSNs play an important role in obtaining the information about the environment and to transfer it to the end user remotely [1]. WSNs are intelligent and automated networks which are used to observe and collect data about the pressure, temperature, vibration, sound, motion, and pollutants and to transfer the data to the main location where it can be analysed [2]. This main location or base station will serve as an intermediate between the user end and the network. The user can get the data from the network by sending queries about the issue to gather the information from the base station. WSNs have a number of nodes which will communicate among

---

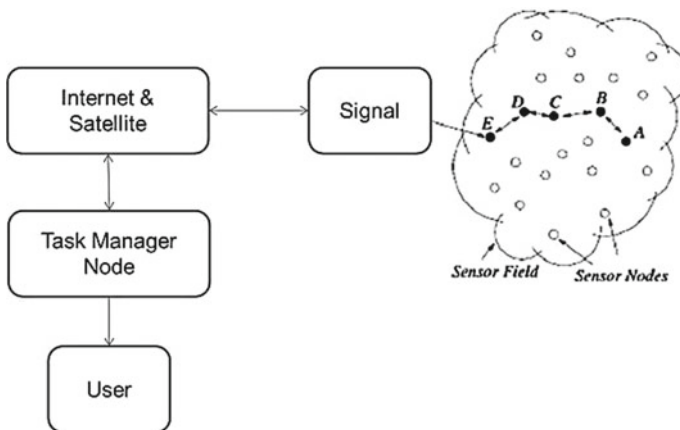
J. N. Swaminathan (✉) · V. Jaikumar · N. S. Kalyan Chakravathy  
QIS College of Engineering and Technology, Ongole 523272, Andhra Pradesh, India

M. Alagumeenaakshi · S. Umamaheswari  
Kumaraguru College of Technology, Coimbatore, Tamil Nadu, India

them with the help of radio signals. In general, the sensor nodes are adept of collecting the data, sensing the required information, processing the data, and communicating it to the networks. These sensor nodes are armed with sensing devices, transceivers, computing devices, and power components. The discrete nodes in WSN have certain limitations such as storage, bandwidth, and processing speed [3]. Once these sensor nodes are installed, they will organise a suitable network infrastructure with multi-point communication. Then, it will respond to the queries sent by the main hub and collect and transfer the information accordingly. Sensors nodes work on either of two modes: event-driven based on the queries from the main network or continuous mode. For processing positioning and location information, these networks use global or local positioning algorithms. Some of the networks are installed with actuators, and they are known as wireless sensor and actuator networks [4]. Figure 1 shows the general block diagram of WSNs.

WSN has wide range of applications, and even it finds applications in health care. Development in the intelligent sensors and WSN has changed the way of patients being diagnosed. A small sensor patch attached on the human body will monitor and collect all the information about the human body and transfer it to the main hub remotely with the help of WSNs [5]. These sensors networks provides a more easily life for the patients and infirms to diagnose their problems remotely.

A proper balance should be maintained between the data processing and communication capabilities in order to avoid the system complexity and high power consumption. If the system is designed with high complexity, then it is difficult for the data to be routed, and high power will be consumed which results in expensive system. These problems inspire the researchers to start working on these applications to overcome these issues [6]. Lot of algorithms and protocols such as energy efficient algorithms, cable mode transition (CMT) algorithm, and particle swarm optimization (PSO) algorithm to trace optical sink were reported. A delay-aware collection network for WSN is to reduce delays while collecting the information from the sensors [7, 8].



**Fig. 1** General block diagram of WSNs

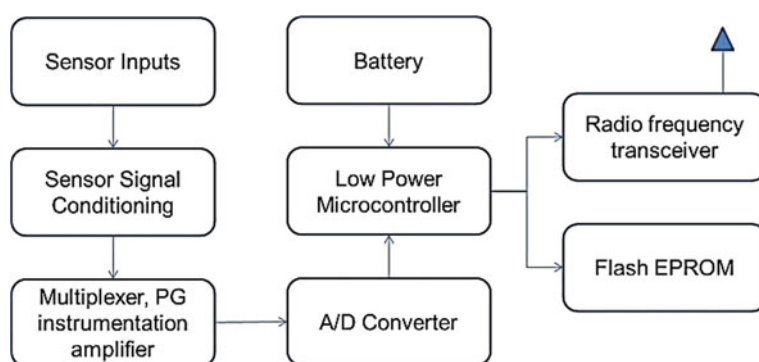


These in turn improve the network lifetime. Also, lot of studies were conducted to increase the network lifetime by using the geometrical clarification to locate the sink. Significant studies were reported on the heterogeneous WSNs in which the nodes have different transmission radii. This architecture overcomes the problems faced in the homogeneous WSNs. In this review, the current achievements of wireless sensor networks in various applications are discussed. Along with that, the security issues and the solutions to overcome those were also discussed briefly.

## 2 Cluster Architecture

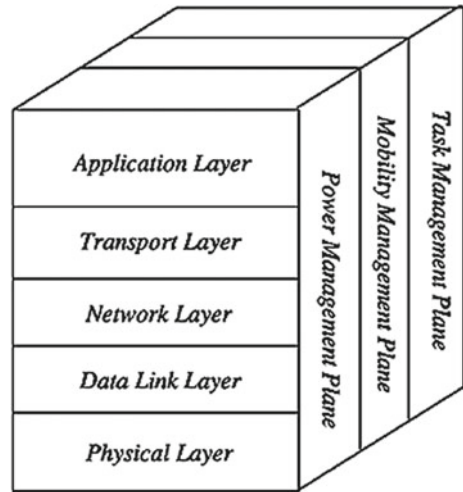
Sensors nodes generally have talent of acquiring information, route the packets to the sink wirelessly, and transfer it to the user end. Figure 2 shows the working block diagram of the nodes. These packets were routed back to the user end based on the protocol stack. The protocol stack is sole responsible for the gathering the packets and routing the packets towards the user end with the help of nodes through the wireless medium.

The protocol stack has equipped with power management plane, computing software, data hyperlink, mobility and undertaking management aircraft, and physical layer. Based on the queries from the network, the application software is deployed and used on the software layer. The delivery layer and network layer perform routing function to transfer and delivery the data exactly to the end point. Avoiding collision and to be conscious about the power, MAC protocol is employed. The basic communication process such as encryption, modulation, and data communication will be carried out by bodily layer. In addition to this, sensor assignment and power consumption reduction will be managed using the energy, venture, and mobility management plane. Figure 3 shows the protocol layers of WSNs.



**Fig. 2** Functional block diagram of sensor nodes

**Fig. 3** Protocol layers of WSNs



### 3 Protocol and Algorithm of WSNs

The main work of the WSNs is to route the data from the particular node to the user end. To achieve the efficient and shortest route path to the base station from node, various routing protocols have been reported [9]. These routing algorithms should be designed to reduce power needs, minimise the packet loss and transfer delay, and to accommodate the time-changing quality of the channel. To overcome these limitations, various routing protocols and strategies have been reported [4, 10–16]. Different classes of routing protocols were reported. In first class, all the nodes were considered as peer nodes which use flat network architecture. It has numerous advantages. Added to that, it also reduces overhead to increase the stability of the infrastructure and reduces the fault tolerance between the nodes. In second class, the nodes are assembled to form clusters. To organise and to perform the function, each individual cluster has a head which has high residual energy. These protocols inflict a network model to attain stability, scalability, and high energy efficiency. The cluster head will take the incharge of organising the activities among the clusters and routing the information among them. Use of clusters results in minimised energy consumption and increased network lifetime. A data-centric approach is used in the third class routing protocols. It is based on the attribute naming in which main node sends attribute-based queries rather than individual node. By allocating the tasks to the sensor nodes and sending queries based on particular attribute, dissemination is achieved. Strategies such as multicasting based on attribute, broadcasting, and geo-casting were used to communicate within the nodes. Location-based routing protocol will come under the fourth class routing protocols. The main node will use the location to sense the sensor node. Apart from these routing protocols, still some of the important routing protocols and algorithms were reported on focusing the energy conservation problems.

### **3.1 Flooding**

It is a general technique which is used most often for discovering the path and information dissemination in wired as well as wireless sensor networks. The strategy used in the flooding technique is simple and does not require high expensive network maintenance and difficult routing algorithm [11]. In this technique, individual node getting a data or packets will route to its neighbour using reactive approach. The packet will follow all the possible paths available to reach the destination. Even if the network architecture alters, the sent packets will track the new path to reach the destination. The packets are replicated indefinitely by the nodes in the networks. Though all the packets were delivered to the destination, the use of high bandwidth due to replicate packets results in costly network.

### **3.2 Gossiping**

To overcome the limitations of flooding, gossiping routing algorithm is proposed in 2002 [17]. The gossiping routing algorithm uses a simple routing rule as like as flooding, but it does not require high expensive topology maintenance or complex routing algorithms. In contrast to the flooding, in gossiping, a data packet is routed to the randomly chosen neighbour instead of all neighbours. The neighbour node once receives the packet and route packets to the next randomly chosen neighbour. At the end, the packet reaches its endpoint by routing packets between the neighbour nodes. Using this, the bandwidth of the network path can be minimised.

### **3.3 Sensor Protocols for Information via Negotiation (SPIN)**

SPIN is protocol that works on the data-centric negotiation-based dissemination for WSNs [18]. The ultimate theme is to proficiently broadcast the opinions collected from the individual sensor nodes to all the sensor nodes. In this technique, each sensor node once receives the packet, and it will advertise the details of the packet among its neighbour nodes. Then, the neighbour nodes will get the packet by sending the request to that node. These techniques overcome the limitations of flooding and gossiping and are considered as energy efficient techniques. SPIN used the format as ADV message, REQ message, and DATA message. ADV is used to advertise, REQ is used to request, and DATA is used to carry the data.

### ***3.4 Low-Energy Adaptive Clustering Hierarchy (LEACH)***

LEACH is a routing protocol which most of the data packets were aimed to receive and transfer the data to the cluster heads rather than nodes [19]. The cluster heads will be responsible for the separating and compressing the data and transmit it to the sink. The cluster head among the nodes will be determined based on the stochastic algorithm. The main motive of this technique is to reduce the energy consumption and to extend the lifetime of the network. It also groups the data to avoid the number of messages. The cluster head will perform different tasks in routing the packets. The cluster will collect the data from the nodes and transmit it to the base station. The next task is to create a TDMA-based schedule for the packets to route it to the respective nodes.

### ***3.5 Threshold-Sensitive Energy Efficient Protocols (TEEN and APTEEN)***

Manjeshwar et al. reported the TEEN and APTEEN routing protocols in 2001 and 2002, respectively [20, 21]. In this technique, cluster head will send a hard threshold to the sensor nodes to sense the required attribute, and for small changes, it will send a soft threshold which will trigger the transmitter to send the packets. Use of this hard threshold results in reduced number of transmission. The nodes will communicate only when the attribute is within the limit of that particular node. Soft threshold limits the transmission as it will transmit only when there is little change. A trade-off can be followed between the data accuracy and energy efficiency. However, these techniques have limitations, and if the nodes do not receive the threshold, it will not communicate which will result in no data transmission to the user.

### ***3.6 Power-Efficient Gathering in Sensor Information Systems (PEGASIS)***

Lindsey et al. proposed a PEGASIS, a group of information gathering and routing protocols for WSNs [22]. Use of PEGASIS enhances the lifetime of the network by using energy efficient techniques and minimises the delay in the data sent to the sink. The same authors have proposed the hierarchical PEGASIS for the WSNs. In PEGASIS, the nodes are assumed to be homogenous, and they have complete idea about the nodes present in the network which are employed with the radio transceiver based on CDMA technique. The main theme of this technique is to develop a routing model to minimise the energy consumption and delay. PEGASIS follows chain structure instead of using other structures like tree- or cluster-based hierarchical.

### **3.7 *Directed Diffusion and Geographic Adaptive Fidelity (GAF)***

A routing protocol for gathering information and dissemination in wireless sensor networks was proposed and named as directed diffusion [23]. In this technique, nodes in the network will interact among them by exchanging messages in a localised interaction with limited network area. The use of this technique will result in the substantial amount of energy savings. The main theme of this technique is to significantly enhance the network lifetime by reducing the energy consumption. Xu et al. in 2001 reported GAF algorithm works on the basis of energy-aware, location-based which can be effectively used for the mobile networks and also be used for WSNs [24]. In this technique, the total network is split into fixed zones, and within each zone, the nodes will collaborate with other nodes. If the node wants to sleep for particular time, then it will assign the other node to perform its role until it is awake. The assigned node will take the full responsibility to monitor and transmit data to the sink on behalf of that node. Because of this, energy consumption can be minimised which will enhance the lifetime of the network.

## **4 Sensor Protocol Analysis**

Several routing protocols and algorithms have been reported especially for WSNs mainly to minimise the energy consumption, improve the stability of the network, and reduce the delay and packet loss. The above-mentioned techniques have their particular merits and demerits on their basis. For example, the flooding and gossiping ensure the delivery of the packets to their user end with the expensive cost of high energy consumption and high bandwidth. Whereas SPIN and LEACH overcome the limitations of flooding and gossiping by using the message request and TDMA-based slot transmission. TEEN and APTEEN communicate only when they receive the request from the user end, but it will also result in loss of communication when they missed the request. PEGASIS minimise the delay and enhance the network lifetime by using CDMA and following chain architecture. Directed diffusion and GAF also proposed to minimise the energy consumption and to enhance the network lifetime. Though various techniques have been proposed, still there is a gap to improve the efficiency of the network, reduce the bandwidth, and to minimise the delay.

## **5 Conclusion**

Wireless sensor networks play a major role in our day-to-day life. It contributes its part in various applications including civilian and medical applications. In this review, the WSN and their architecture were discussed in detail. The routing protocol and

algorithm especially reported for wireless sensor network were also discussed in this review. The issues faced by the several routing algorithms based on energy efficient, network stability, and minimal delay were discussed for each algorithm. However, still, these factors are to be addressed in order to design a high efficient low-power wireless sensor networks. The future research should be engaged in increasing the throughput of the WSN networks. The research in WSN will advance the technology development particularly in smart home environment. It also contributes in medical field in which the physiological signals can be sensed from the human body and transmit it to the user end using wireless sensor networks. Use of WSNs will improve the living style of humans in near future.

## References

1. Kandris D, Nakas C, Vomvas D, Koulouras G (2020) Applications of wireless sensor networks: an up-to-date survey. *Appl Syst Innov*
2. Srivastava S, Singh M, Gupta S (2018) Wireless sensor network: a survey. In: 2018 International conference on automation and computational engineering, ICACE 2018
3. Kobo HI, Abu-Mahfouz AM, Hancke GP (2017) A survey on software-defined wireless sensor networks: challenges and design requirements. *IEEE Access*
4. Akkaya K, Younis M (2005) A survey on routing protocols for wireless sensor networks. *Ad Hoc Netw*
5. Dhanabalan SS, Sriram S, Walia S, Avanimathan SR, Carrasco MF, Bhaskaran M (2021) Wearable Label-Free Optical Biotectors: Progress and Perspectives,” *Adv. Photonics Res.*, 2021.
6. Buratti C, Conti A, Dardari D, Verdona R (2009) An overview on wireless sensor networks technology and evolution. *Sensors*
7. Cheng CT, Tse CK, Lau FCM (2011) A delay-aware data collection network structure for wireless sensor networks. *IEEE Sens J*
8. Cheng CT, Tse CK (2013) A delay-aware network structure for wireless sensor networks with consecutive data collection processes. *IEEE Sens J*
9. Sharma S, Jena SK (2011) A survey on secure hierarchical routing protocols in wireless sensor networks. In: *ACM international conference proceeding series*
10. Singla P, Munjal A (2020) Topology control algorithms for wireless sensor networks: a review. *Wirel Personal Commun*
11. Akyildiz IF, Su W, Sankarasubramanian Y, Cayirci E (2002) Wireless sensor networks: a survey. *Comput Netw*
12. Akyildiz IF, Su W, Sankarasubramanian Y, Cayirci E (2002) A survey on sensor networks. *IEEE Commun Mag*
13. Boukerche A (2008) Algorithms and protocols for wireless sensor networks
14. Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. *IEEE Wirel Commun*
15. Santi P (2005) Topology control in wireless ad hoc and sensor networks. *ACM Comput Surv*
16. Waharte S, Boutaba R, Iraqi Y, Ishibashi B (2006) Routing protocols in wireless mesh networks: challenges and design considerations. *Multimed Tools Appl*
17. Braginsky D, Estrin D (2002) Rumor routing algorithm for sensor networks. In: *Proceedings of the ACM international workshop on wireless sensor networks and applications*
18. Kulik J, Heinzelman W, Balakrishnan H (2002) Negotiation-based protocols for disseminating information in wireless sensor networks. *Wirel Netw*

19. Bagci F (2016) Energy-efficient communication protocol for wireless sensor networks. *Ad-Hoc Sens Wirel Netw*
20. Manjeshwar A (2002) APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. *ipdps ipdps*
21. Manjeshwar A, Agrawal DP (2002) APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless. In: *Proceedings—international parallel and distributed processing symposium, IPDPS 2002*
22. Lindsey S, Raghavendra CS (2002) PEGASIS: Power-efficient gathering in sensor information systems. In: *IEEE aerospace conference proceedings*
23. Intanagonwiwat C, Govindan R, Estrin D (2000) Directed diffusion: a scalable and robust communication paradigm for sensor networks. In: *Proceedings of the annual international conference on mobile computing and networking, MOBICOM*
24. Xu Y, Heidemann J, Estrin D (2001) Geography-informed energy conservation for ad hoc routing. In: *Proceedings of the annual international conference on mobile computing and networking, MOBICOM*

# Building a Model of Signature Verification and Identification System Based on Deep CNN (ConvNet)



R. Shirisha, P. Divya, V. B. Aiswarya, S. Yoshitha, Jasim Sadique, and Anitha Premkumar

**Abstract** Signatures are commonly used as a unique way of identifying and verifying a person's identity. Legal documents like bank cheques and transactions, government documents, and affidavits require signature-based verification to identify the authenticity of the person. The signature identification and verification system help to distinguish whether an input signature is honest or a fake. For very long years, this was considered to be a challenging mission when the offline mode was taken into consideration. The offline mode uses scanned signature images, where the online mode of signing process isn't existing. To overcome the signature verification challenge, an application of deep learning techniques to identify feature representations from signature images can be leveraged. In this paper, we have demonstrated how to identify valid and invalid signatures by building a signature verification and identification system that uses deep convolution neural network (ConvNet).

**Keywords** Deep learning · Convolution neural network (ConvNet) · Signature verification

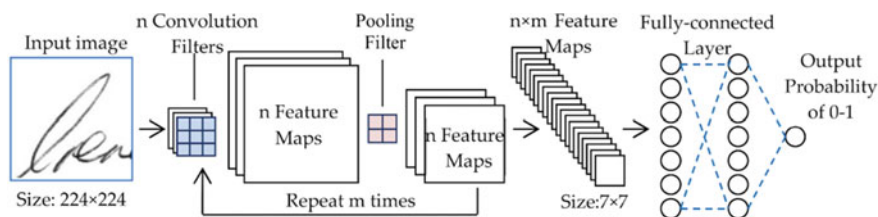
## 1 Introduction

A signature is an individual's name written in a particular manner as a type of identification in approving checks, reports, letters, visa, passport, etc. Signatures are the most legally accepted means for individual verification. Therefore, it is a modality with a high level of attacks like forgeries on documents. For instance, whenever a person deposits a cheque through the banking system, it will verify the signature of the person against the record of signatures stored at bank database before proceeding with the cheque transaction. Every year, worldwide many incidents are taking place on signature forgery. Here are some recent forgery cases which took place in India.

---

R. Shirisha · P. Divya · V. B. Aiswarya · S. Yoshitha · J. Sadique · A. Premkumar (✉)  
Department of Computer Science, Presidency University, Bengaluru, Karnataka, India  
e-mail: [anithapremkumar@presidencyuniversity.in](mailto:anithapremkumar@presidencyuniversity.in)





**Fig. 1** Convolution neural network (CNN) architecture

- As per the recent report of NDTV, the Mumbai crime branch arrested Ace Indian vehicle architect person in a case regarding forgery and cheating, reportedly in a racket which involves copying vehicle enrollments.
- According to Times of India, an offense was registered against three people, and a lady was also involved on January 21 [2021] for allegedly forging a birth certificate and presenting it to the city enterprise (BNMC) in Maharashtra.
- According to the report from Times of India, on March 11, 2021, in Ahmedabad, Santej police of Gandhi Nagar arrested a person who is managing director of private limited, where a complaint was registered by his elder brother alleging that the person availed Rupees 200 crore loan by forgery.

In these scenarios, a systematic approach for authenticating the signature is necessary to stop such forgery. This paper helps us to improve the verification of person's signature using deep convolution neural network algorithm. A convolutional neural network (ConvNet) is a part of deep neural network which has one or many layers of convolution as shown in Fig. 1. The main purpose of CNN [11–13] is for image processing, segmentation, classification, and also for other autocorrelated data. The term convolution in convolution neural network usually indicates the mathematical function of convolution which is an uncommon sort of linear operation where two functions are increased to create a third function.

- **Convolution Layer**—This is the primary layer that is used to extract various features from the input signature image. The resulting output is named as a feature map which gives data about the image which includes edges and corners.
- **Pooling Layer**—After the convolution layer completes the process, the pooling layer takes place. This layer reduces the size of the convolved feature map to decline the computational expenses. This is performed by reducing the links among layers and individually operates on each featuring map. The pooling layer typically fills in as a scaffold between the convolutional layer and the FC layer.
- **Fully connected Layer**—The fully connected layer is used to connect neurons between two various layers, and it also consists of weights and biases. These layers are generally located before the output layer and structure the last couple of layers of a CNN architecture.

This paper is organized into sections as follows: In Sect. 2, we discuss the paper motivation. In Sect. 3, we presented the related works carried out in this field. A

system design is explained in Sect. 4. We analyze the result of a system in Sect. 5. In Sect. 6, we conclude the paper.

## 2 Motivation

The signature verification task is extremely basic and frequently presents challenges like high fluctuation. For example, an individual's signature may differ each time and may differ totally with age, and also, signature may differ according to their mood, likeliness between signatures of various individuals and similitude in duplication or falsification of one's signature. This project focuses on handwritten signature verification using convolution neural networks for training the dataset of digitally scanned handwritten signatures and showing the accurate result of the signature whose legitimacy is needed to be determined. The person who verifies signatures from documents is taking a huge risk as it will be his/her final decision to conclude that signatures on documents are genuine. This kind of validation process is risky, and if any inaccurate authentication result may lead to serious damage and consequences. The shape varieties are normal, and these are endeavored at catching their belongings in type of CNN highlights which are classified as genuine or forged. The main motivation of this paper is to build a model of a system which can help to verify and identify the signature whether it is valid or invalid and also improve the authenticity of signature using convolution neural network algorithm.

## 3 Related Work

There have been various studies and research in this field carried out to deal with signature verification, both online and offline; each approach has its own upsides and downsides. Table 1 shows that the list of related works carried out in this field.

## 4 System Design

### 4.1 Building Components of a System

In Fig. 2, the building components of signature verification and identification system are shown. The images of different signatures are stored in a file and feed as input to a system. The ConvNet algorithm is implemented using Python language which uses Keras with TensorFlow as backend to learn the samples of the input signatures. Keras is a high-level application programming interface (API) which is a wrapper to the backend libraries of TensorFlow. The details of the components used for building a system are shown in Table 2.

**Table 1** List of related works carried out in this field

Author	Title	Year	Publication	Method used
Abughfa and Elmadani [1]	Offline signature verification based on image processing and Hu moment	2016	International journal of scientific research in network security and communication	– Statistical parameter -mean squared Error and Hu moment
Kao and Wen [2]	An offline signature verification and forgery Detection method based on a single known Sample and an explainable deep learning approach	2020	Applied sciences	– CNN – Deep learning, forensic science
Engin et al. [3]	Offline signature verification on real-world documents	2020	IEEE/conference on computer vision and pattern recognition workshops	– A stamp cleaning method based on CycleGAN – Signature representation based on CNNs
Navid et al. [4]	Signature verification using convolutional neural network	2019	IEEE international conference on robotics, automation, artificial intelligence and Internet of things (RAAICON)	VGG-19
Gopichand et al. [5]	Digital signature verification using artificial neural networks	2019	International journal of recent technology and engineering (IJRTE)	Deep learning, digital configured replica, forgery, signature
Yapic et al. [6]	Convolutional neural network-based offline signature verification application	2018	International congress on big data, deep learning, and fighting cyber terrorism	Deep learning, convolution neural network
Suryani et al. [7]	Offline signature recognition and verification system using efficient fuzzy Kohonen clustering network (EFKCN) algorithm	2017	International conference on computer science and computational intelligence, (ICCCSCI)	Efficient fuzzy Kohonen clustering network (EFKCN) algorithm

(continued)

Table 1 (continued)

Author	Title	Year	Publication	Method used
Shrestha et al. [8]	Offline signature verification using convolutional neural network	2017	Kathford international college of engineering and management	– Convolutional neural network machine learning techniques – ML, ANN
Sabbagh et al. [9]	Signature verification	2016	Research gate publications-Technical Report	Artificial neural network
Julita et al. [10]	Online signature verification system	2015	Research gate publication (international colloquium on signal processing and its applications (CSPA))	Support vector machine

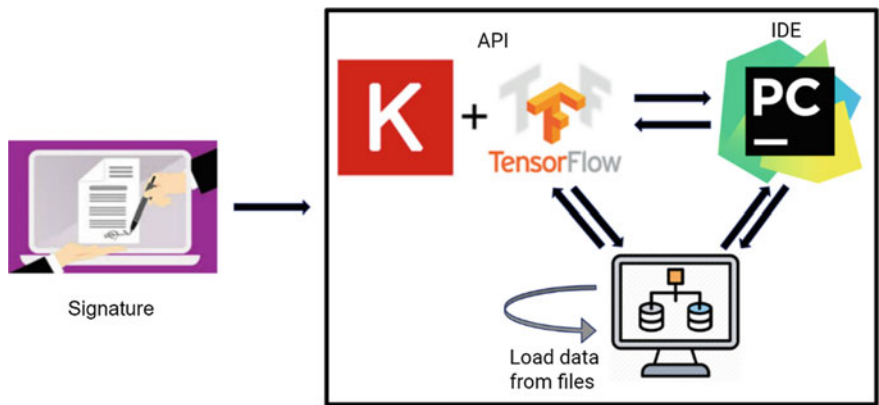


Fig. 2 A building components of a Signature Verification and Identification System

Table 2 Development environment of a signature verification and identification system

Components	Version	Description
PyCharm	2020.2.1*64	Python IDE
TensorFlow	2.5.0-rc3	Open-source library
Python	3.9.0	Programming language
Keras	2.5.0	High level API for TensorFlow
Flask		Web application framework
HTML		Programming language for frontend design

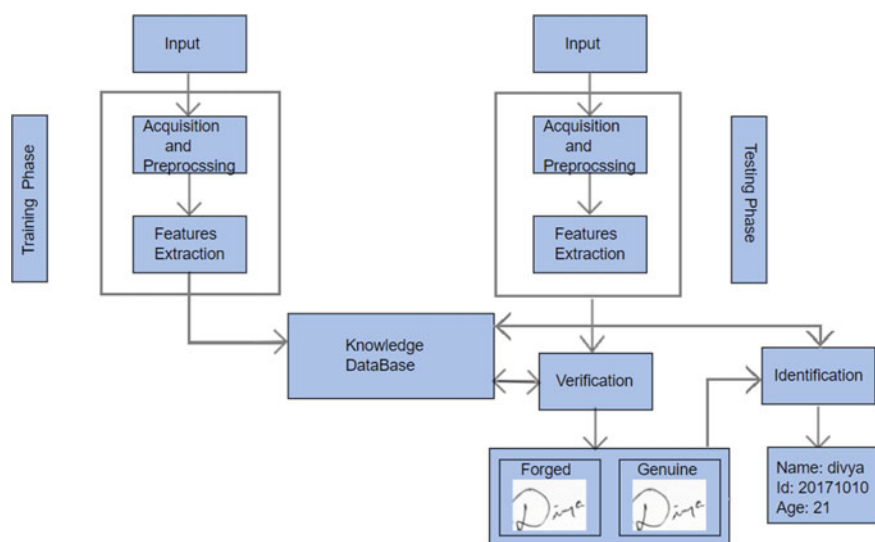


Fig. 3 A workflow of a signature verification and identification system

## 4.2 Workflow of a System

In Fig. 3, working model of signature verification and identification system is shown. This model consists of two different phases, namely testing phase and training phase. The preprocessed signature images are loaded into the system where system will be trained on those images to value the performance.

- **Data Acquisition**—An input signature will be collected by data acquisition. This will collect the digital image format of signature from handwritten signatures.
- **Preprocessing**—In this process, all the digital image format of signature is normalized, resized to proper dimensions, background sounds or noise will be removed, and it will resize the image signature according to the template which will be further used for extracting the features of the signature image.
- **Feature Extraction**—To perform the feature extraction process, a convolution neural network algorithm has been used. During the feature extraction process, the system examines a given pattern and records the certain features, in order to submit the structured data in the form of an observation sequence. And also, this process recognizes and discriminates a person's signature from one another. A feature extraction process plays an important role to improve the accuracy of a system.
- **Knowledge database**—Once the features are extracted, the training dataset will be stored in the knowledge database. This knowledge database will be used in further steps of the verification process to verify a signature.
- **Verification process**—After the feature extraction process, the testing image of signature will be verified. Finally, the testing image of the signature will be

compared with the signature images stored in the knowledge database and gives the result whether the signature is valid or in-valid.

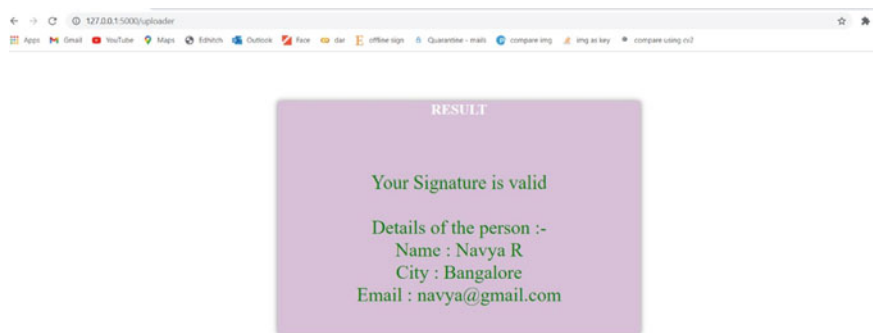
- **Identification process**—At the end of the verification process, if the result is valid, then the signature image shows the identification details of that particular signature and this identification details includes name, id, and age. If the verification result is forged, then no identification details will be displayed.

## 5 Implementation Result

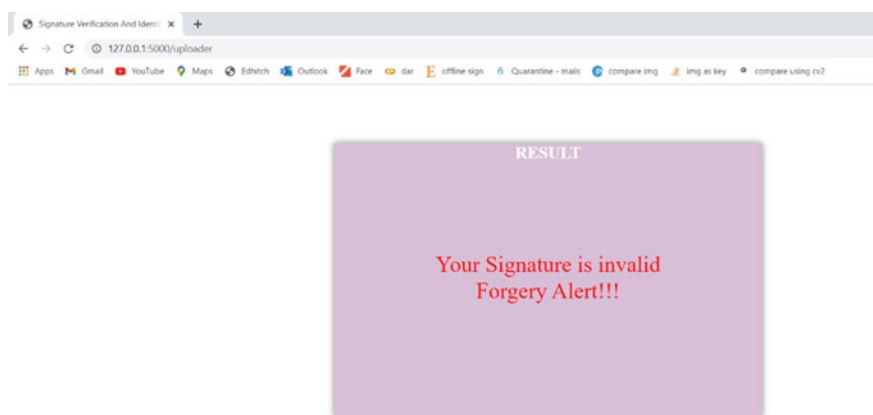
The handwritten user’s signature is scanned and uploaded into the system via graphical user interface (GUI) as shown in Fig. 4. This uploaded signature is compared with the original image present in the system and gets verified. Once the verification process is completed, it retrieves the details of the matched signature and result is shown in Fig. 5. If it is not matched with the signature, a system gives forgery alert as shown in Fig. 6. Hence, a system will be highly secured and encrypted for a person to access. As the system is using TensorFlow as backend and Keras a wrapper to it, this is secured from forgery because it requires the genuine signature and avoids forgery due to multiple iterations of signature check and the probabilities obtained will be matched.

**Fig. 4** GUI of signature verification and identification system





**Fig. 5** Retrieves the details of the matched signature after the verification process completes



**Fig. 6** Forgery alert by system

## 6 Conclusion

This paper demonstrated a model of signature verification and identification system using a recent technology called deep convolution neural network to tackle the signature forgery problem. This model helps us to differentiate genuine signatures from forged ones, and it helps in declaring a legal ownership of an asset. This model uses different features extracted from preprocessed signature images. Compared to the previous works, this model is a unique design for retrieving the details of a person which is highly encrypted for accessing and probability is high to avoid forgery. As this model is focused on deep convolutional neural networks, the performance of the neural network can be expected to be better but time for training and testing may increase.

## References

1. Abughfa AKI, Elmadani AB (2016) Offline signature verification based on image processing and hu moment. *J Sci Res Netw Secur Commun* 4(5). Alfateh University, Tripoli, Libya
2. Kao H-H, Wen C-Y (2020) An offline signature verification and forgery detection method based on a single known sample and an explainable deep learning approach. *Appl Sci*. Department of Forensic Science, Central Police University, Taoyuan 33304, Taiwan
3. Engin D, Kantaraci A, Arslan S, Kermal H (2020) Offline signature verification on real-world documents. In: *IEEE/Conference on Computer Vision and Pattern Recognition workshops*, June 2020, pp 808–809. Yapi Kredi Technology, Istanbul Technical University
4. Navid SAM, Priya SH, Khandakar NH, Ferdous Z, Haque AB (2019) Signature verification using convolutional neural network. In: *IEEE international conference on robotics, automation, artificial intelligence and internet-of-things (RAAICON)*, Dhaka, Bangladesh
5. Gopichand G, Sailaja G, Vinod Kumar VN, Samatha T (2019) Digital signature verification using artificial neural networks. *Int J Recent Technol Eng (IJRTE)* Blue Eyes Intell Eng 7(552)
6. Yapic MM, Tekerek A, Topaloglu N (2018) Convolutional neural network based offline signature verification application. In: *International congress on big data, deep learning and fighting cyber terrorism*, Ankara, Turkey, 3–4 December, 2018
7. Suryani D, Irwansyah E, Chindra R (2017) Offline signature recognition and verification system using efficient fuzzy Kohonen clustering network (EFKCN) algorithm. In: *International conference on computer science and computational intelligence, (ICCCSI)*, Bali, Indonesia, 13–14 October 2017
8. Shrestha N, Sharma P, Ghimire, Neupane S, Pokharel S (2017) A project report on “offline signature verification using convolutional neural network”. Kathford International College of Engineering and management, Balkumari, Lalitpur, Aug 2017
9. Sabbagh SA, Shanwar B, Zarka N (2016) A technical report on “signature verification. Researchgate publications-, Higher Institute for Applied Sciences and Technology Communications Department, June 2016
10. Julita, Fauziyah S, Azlina O, Mardiana B, Hazura H, Zahariah AM (2015) Online signature verification system. Researchgate publication, International Colloquium on Signal Processing & Its Applications (CSPA), Ayer Keroh, Melaka
11. Lin G, Shen W (2018) Research on convolutional neural network based on improved Relu piecewise activation function. *Sci Direct* 131:977–984. *ICICT-2018*
12. Jaswal D, Sowmya V, Soman KP (2014) Image classification using convolution neural networks. *Int J Adv Res Technol* 3(6)
13. Liu G, Zhang J (2020) CNID: research of neural intrusion detection based on convolution neural network. *Discrete Dyn Nat Soc* 2020:11. Article ID 4705982. Hindawi



# Product Review Analysis



L. V. Krishna Rao, S. Sai Vamsi, Sk. Sabiha Naazneen, S. Gowthami,  
and G. Bindu Bharati

**Abstract** Many websites sell products and customers buy these products. Customers give their rating to the products they bought from these websites which helps genuine rating of the products. But some websites for fame and profit give fake reviews so it is hard for a customer to know which product is genuine. So, because of this, we detect fake reviews and genuine reviews using sentimental analysis and here we calculate the sentimental score and remove negative words in order to detect fake and genuine reviews. After detecting fake and genuine reviews using with stop words and without stop words, we find the accurate algorithm to find these fake and genuine reviews and calculate the accuracy here. Later, we choose the best accurate algorithm. We plot the visualization graphs for fake and genuine reviews.

**Keywords** Sentiment analysis · Naive bayes · K Nearest neighbor · Decision tree · Support vector machine

## 1 Introduction

Sentimental analysis is used to classify the reviews as fake or genuine. To know whether the product is genuine or fake we use this sentimental analysis technique to find the fake and genuine reviews which also helps the customer [1]. Detecting the fake and genuine reviews is quite challenging here we use this sentimental analysis techniques we take the reviews and calculate the sentimental score of each review and then based on that we detect the review as fake positive, fake negative, true positive, true negative [2]. Here, it will find the polarity of each word in the review. Here, we take the dataset, and we apply the four supervised algorithms which are mainly used for classification purpose. Here, we used the algorithms like SVM, KNN, DT-J48, NB. We used with using the stop words and without using the stop words for all the four algorithms, and we calculate the accuracy here for each algorithm and find

---

L. V. Krishna Rao (✉) · S. Sai Vamsi · Sk. Sabiha Naazneen · S. Gowthami · G. Bindu Bharati  
JNTU Kakinada, Lakireddy Bali Reddy College of Engineering, Mylavaram, India

**Table 1** Study of types of algorithms

Algorithm	True positive		False positive		Weighted average TP	Weighted average FP
	Normal	Anomaly	Normal	Anomaly		
SVM	0.972	0.964	0.036	0.028	0.968	0.032
ANN	0.960	0.959	0.040	0.041	0.960	0.040
KNN	0.986	0.990	0.014	0.010	0.988	0.012
NB	0.911	0.887	0.091	0.123	0.899	0.101
DT	0.995	0.990	0.004	0.010	0.993	0.007

out the best accurate algorithm in both the cases with using stop words and without using stop words. Later, we find out the fake and genuine reviews as fake positive, fake negative, true positive, true negative which helps in classify the product as fake or genuine, in this we found that SVM with using stop words and without using stop words is more accurate than other algorithms like DT-J48, KNN, NB. The main goal is to classify the reviews as fake or real and find the best accurate algorithm, while classifying this. The rest of the rest of the paper consists of methodology, related work, experimental results, and conclusion. We detect the fake reviews for the products with the help of sentimental analysis [3].

## 2 Related Works

In this section, we have methods to evaluate the performance in order to detect the reviews as fake and real also to find the accuracy (Table 1).

The dataset consists of reviews in the textual form as well as rating from 1 to 5 [4]. Here, the data we collected is further processed in order to classify the review as fake or real. We have used the four supervised machine learning algorithms DT-J48, SVM, KNN, and NB to classify the reviews of products as fake or real [5]. In this paper, the results are shown that SVM is more accurate than other algorithms in both cases with and with stop words [6].

## 3 Methodology

In order to test our methods for classification of reviews, the collected data set is assigned a sentiment value for analyzing the reviews more accurately.

### Step 3.1: Review Data Set

Here, we have taken the amazon data set in excel format and it consists of rows and columns. The data set must be cleaned in order to apply the algorithms, so we first apply the preprocessing steps to the taken dataset in order to clean it and use it [7].

The data set is collected to analyze the sentiment value for detecting the reviews as fake or genuine.

### **Step3.2: Data Preprocessing**

In this step, we have the data preprocessing methods in order to clean the data. We have the following steps to clean our dataset of reviews.

#### **3.1 Punctuation**

In the textual dataset, we have taken has many punctuations. These are not necessary as it does not add any meaning. The string has around 32 punctuations. These are given below: `'";}{[N_-!@#$$%^&* > < ?" ~ ' (Fig. 1).`

#### **3.2 Tokenization**

Tokenization is the way to separate the textual data into small units like words, characters, or substrings. Here, we create a vocabulary of words in this step. It is the process of replacing the given data into something without hindering the security. The regular expression is a tool is used to match the similar patterns, with the help regular expression, the sentence is tokenized.

##### **Stop words**

Stop words are non-useful words which are found in the text they do not help in classifying the review as positive or negative. Here, we use “nltk” library for stop words. We have stop words in this library like below. In this library around 179, stop words are there. [“you”, “me”, “I”, “us”, “the”, “he”, “she”, “it”, “they”]

##### **Lemmatize/Stem**

Lemmatization is used to minimize the variations of the given word. If you have a review given by customer, they use different words like “walking” its basic form is “walk” thus it select them. It will help in selecting these correct lemmas based up on the context.

##### **Other steps**

In this step, we have URLs or html tags or emojis or numbers these can be re moved in this step.

### **Step 3.4: Sentiment Classification Algorithms**

Here in this step, we have used the classification algorithms. There are many but here we used the four main algorithms which are SVM, KNN, DT-J48, NB.

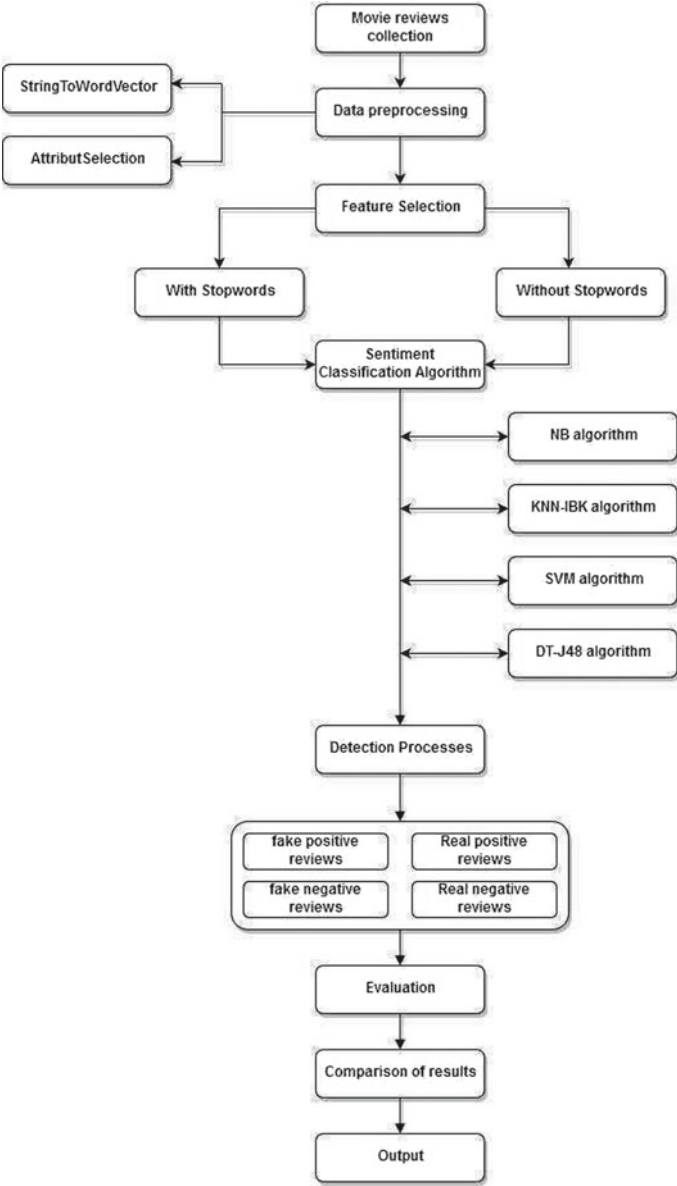


Fig. 1 Flow diagram

**Naive Bayer’s (Nb)**

This is the supervised learning algorithm, which is used for solving classification problems. This algorithm is one of the simple and effective classification algorithms. It is the best choice for text classification algorithms [8].

**Support Vector Machine (SVM)**

This is the one of the most popular supervised machine learning algorithms which can be for classification as well as regression [9]. The goal SVM is used to create a best line which helps in creating hyperplane. The extreme cases are called support vectors.

**KNN**

This algorithm is simple it takes k data points and later check for the similarity of points and finds the smallest distance between the points in order to find the which cluster the data point belongs thus it works [10]. It is more effective if the dataset that is training dataset is large.

**Decision Tree**

Decision tree algorithm is a supervised learning technique that can be used for solving classification problems [11]. It looks like a tree shape and the inner nodes are used to represent the features in the taken data set and also the branches define the decision rules, and the leaf nodes define the outputs.

**Step 3.5: Detection Process**

After the step 4, we have to predict the output and generate the confusion matrix that classifies the review as positive or negative. We classify the reviews as follows:

True Positive (TP): In this case, if the value in predicted case and actual case is same then we say true positive.

True Negative (TN): In this case, the predicted value matches the actual value. The actual value was negative, and the model predicated a negative value.

False Positive (FP): In this case, the actual value was negative, but the model predicted a positive value.

False Negative (FN): In this case, the predicted value was falsely predicted. The actual value was positive, but the model predicted a negative value. Various classification algorithms are there for detecting fake positive and negative reviews (Table 2).

The above table provides the performance of the classification algorithms by making predictions on the test data and accuracy of our classification model. The

**Table 2** Performance of classification algorithms

Total Predictions (n)	Actual: no	Actual: yes
Predicted: no	True negative	False positive
Predicted: yes	False positive	True positive

confusion matrix is used to find the accuracy as well as precision using the below formulas:

$$\text{Fake Positive} = \text{"FP/FP + TN"}$$

$$\text{Fake Negative} = \text{"FN/TP + FN"}$$

$$\text{Real Positive} = \text{"TP/TP + FN"}$$

$$\text{Real Negative} = \text{"TN/TN + FP"}$$

### Accuracy

It is defined as how often the model gives the correct output. The formula for calculating the accuracy is given below.

$$\text{Accuracy} = \text{TP + TN/TP + FP + FN + TN}$$

Precision: It can be defined as how many of them were actually true. It can be calculated using the below formula:

$$\text{Precision} = \text{TP/TP + FP}$$

For every algorithm, we generate different confusion matrix.

### Step 3.6: Comparison of Results

We have taken the four classification algorithms and generated the different accuracies for the amazon dataset, and we identified the best accurate algorithm to find the reviews as positive or negative.

### Step 3.7: Deletion of Fake Reviews

Here, we delete the fake reviews. Which are detected by using above algorithms. This helps in genuine rating of results, and this reduces the amount of data set as well.

## 4 Experimental Results

Here, we have taken the four classifiers and classifying the reviews as fake or real the below table represent the fake and real reviews for each algorithm. We have taken with stop words as well as without stop words [12].

### Without Stop Words

The below table shows the real and fake reviews for all methods using without stop words (Table 3).

**Table 3** Real and fake reviews for all algorithms without stop words

Classification algorithms	Sentiment anal	Real	Fake
Naïve bayes	Real	1745	617
	Fake	706	2599
SVM	Real	1716	629
	Fake	443	2879
KNN	Real	1738	624
	Fake	713	2592
DT-J48	Real	1735	627
	Fake	710	2595

**Table 4** Classification algorithms and their accuracy

Classification algorithms	Accuracy %
Naïve bayes	76.3
SVM	82.2
KNN	68.3
DT-J48	70.3

The number of real and fake predictions made by the classification model is compared with the actual results in the test data is shown in the confusion matrix. The confusion matrix is obtained after implementing the four algorithms. The confusion matrix generated is given below. It shows the accuracy percentage for each algorithm without stop words (Table 4).

From the above table, the SVM has high accuracy compared to the other classifiers (Table 5 and Fig. 2).

**With Stop Words**

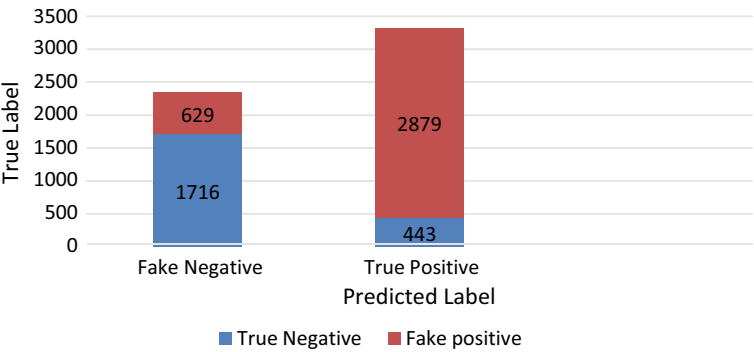
In the above, we compared different algorithms without stop words. In this, here we compare different algorithms using with stop words. Here, the below table represents the confusion matrix created without using stop words (Table 6)

The accuracy table for without using stop words for all the methods are given below (Table 7).

From the above table SVM has high accuracy compared to the other classifiers [12]. Precision, recall, F1 score, and confusion matrix are shown below (Table 8 and Fig. 3).

**Table 5** Precision, recall, F1 score, and confusion matrix

	Precision	Recall	F1-score	Support
Fake	0.79	0.75	0.77	2338
Genuine	0.82	0.86	0.85	3329



**Fig. 2** The predicted label and true labels

**Table 6** The confusion matrix created without using stop words

Classification algorithms	SA	Real	Fake
Naïve bayes	Real	1594	698
	Fake	685	2690
SVM	Real	1686	606
	Fake	437	2938
KNN	Real	1589	703
	Fake	690	2685
DT-J48	Real	1581	711
	Fake	698	2677

**Table 7** The accuracy table for without using stop words for all the methods

Classification algorithms	Accuracy %
Naïve bayes	75.5
SVM	81.5
KNN	68.9
DT-J48	70.3

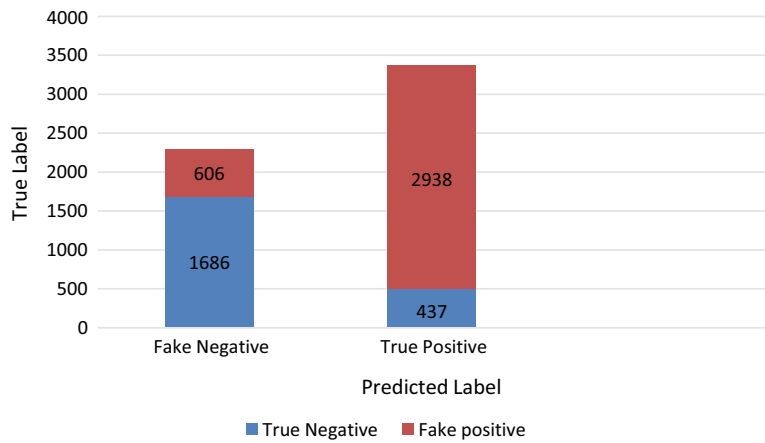
**Table 8** Precision, recall, F1 score, and confusion matrix with stop words

	Precision	Recall	F1-score	Support
Fake	0.79	0.74	0.76	2292
Genuine	0.83	0.87	0.85	3375

**4.1 Evaluation Parameters and Accuracy**

The algorithms which are used out of all the SVM with stop words and without stop words have the high accuracy than any other algorithm we have 82% accuracy for



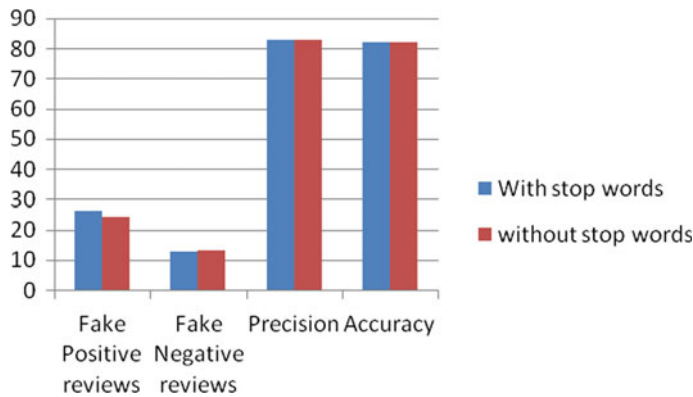


**Fig. 3** The predicted label and true label in with stop words

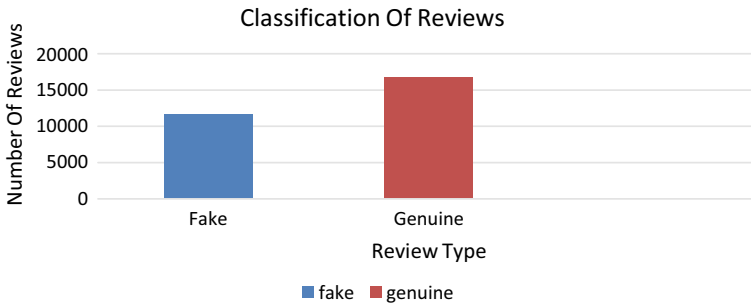
**Table 9** Features and parameters in both cases stop words and without stop of SVM

Features and parameters	With stop words	without stop words
Fake positive reviews	26.4	24.6
Fake negative reviews	12.9	13.6
Precision	83	83
accuracy	82	82

SVM [13] as shown below in both the cases and also, we plotted the visualization graph which is shown (Table 9; Figs. 4 and 5).



**Fig. 4** Graph shows the fake and precision values in both cases



**Fig. 5** Shows the overall fake and genuine reviews

## 5 Conclusion

In this, we have taken the amazon data set in order for classification of reviews as fake or real of products. We have used the four machine learning algorithms which are KNN, NB DT-J48, SVM. Here, we calculated the accuracy of each algorithm by generating the confusion matrix of these four algorithms with and without stop words and out of all SVM is accurate than other three algorithms. In future, we would like to extend the data set that contains image data of products.

## References

1. Liu B (2012) Sentiment analysis and opinion mining. *Synth Lect Hum Lang Technol* 5(1):1–167
2. Malbon J (2013) Taking fake online consumer reviews seriously. *J Consum Policy* 36(2):139–157
3. Elshrif E, Gherbi A (2017) An empirical study on detecting fake reviews using machine learning techniques. In: *The seventh international conference on innovative computing technology (INTECH 2017)*
4. Tian N, Xu Y, Li Y, Abdel-Hafez A, Audun Josang (2014) Generating product feature hierarchy from product reviews. In: *International conference on web information systems and technologies*, pp 264–278. Springer International Publishing
5. Xia R, Zong C, Li S (2011) Ensemble of feature sets and classification algorithms for sentiment classification. *Inf Sci* 181(6):1138–1152
6. Tian N, Xu Y, Li Y, Abdel-Hafez A, Jøsang A (2014) Product feature taxonomy learning based on user reviews. In: *WEBIST no 2*, pp184–192
7. Abdel-Hafez A, Xu Y (2013) A survey of user modelling in social media websites. *Comput Inf Sci* 6(4):59
8. Barbu T (2012) SVM-based human cell detection technique using histograms of oriented gradients. *cell* 4:11
9. Esposito G (2014) LP-type methods for optimal transductive SVMs. Vol 3. Gennaro Esposito, PhD
10. Kalaivani P, Shunmuganathan KL (2013) Sentiment classification of movie reviews by supervised machine learning approaches. *Indian J Comput Sci Eng* 4(4):285–292

11. Pang B, Lee L (2004) A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on association for computational linguistics, p 271. Association for Computational Linguistics
12. Vinodhini G, Chandrasekaran RM (2012) Sentiment analysis and opinion mining: a survey. *Int J* 2(6):282–292
13. Xu G, Cao Y, Zhang Y, Zhang G, Li X, Feng Z (2016) TRM: computing reputation score by mining reviews. In: Workshops at the thirtieth AAAI conference on artificial intelligence

# Optimum Route Computation in a Chaotic Artificial Climbing Wall



Dibyendu Seal and Reshmita Seal

**Abstract** Artificial wall climbing has become very popular now-a-days in the area of adventure sports. Climbing route computation aims to detect the best route suitable for climber. This paper proposes an efficient route finding technique in a climbing wall where the climbing holds are placed randomly. The method analyzes the distances between climbing holds and also considers the ability of climber to shift body to the adjacent hold. This method gives better insight about the climbing route to a route setter and also proposes multiple feasible routes in an artificial climbing wall.

**Keywords** Shortest path · Wall climbing · Rock climbing · Undirected graph · Dijkstra algorithm · Stride · Python · NetworkX · Route setting

## 1 Introduction

In Olympic Games, 2020, with the inclusion of artificial rock climbing the popularity of this sport has been boosted up significantly. This increase in popularity has caused a new demand for route setters [1] to come up different possible and innovative climbing routes. Currently, the setting of a climbing route relies on the creativity and climbing experience of a route setter. The objective of this paper is to find the optimum climbing route in an artificial climbing wall where the climbing holds are distributed randomly, which gives the artificial climbing wall as essence of a natural rock wall. This paper also computes the shortest climbing route according to height of climbers and tells route setters about the various feasible climbing routes on the wall and the difficulty level of the route w.r.t other routes.

---

D. Seal (✉)

Tata Consultancy Services, Kolkata, India

R. Seal

Techno Main Saltlake, Kolkata, India

## 2 Related Work

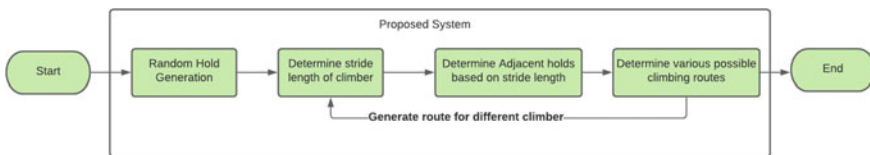
The first indoor climbing wall opened in 1974 in Bolzano, Italy [2]. The problem of setting a good climbing route being done in many ways around the world by various researchers. Frank Stapel tried to generate climbing routes by creating a greedy algorithm using heuristics method. Then, the generated routes were compared with existing Moon-Board routes by experienced climbers. But the assessment did not achieve the benchmark [1]. Kempen developed a machine-learning based model which predicts the difficulty grades of climbing routes. He developed CRDL (climbing route description language) for generating new routes using different types of holds which parsed the climbing route in a sequence of symbols [3]. Naderi et al. proposed a technology to mental practice in sports, which they denoted as computer-aided imagery (CAI). They were presented a novel interactive computational sport simulation for exploring and planning movements and strategies. But their research did not give accurate simulation for both gross-motor body movements and fine-motor control of the myriad ways in which climbers can grasp climbing holds or shapes. [4].

## 3 Proposed Work

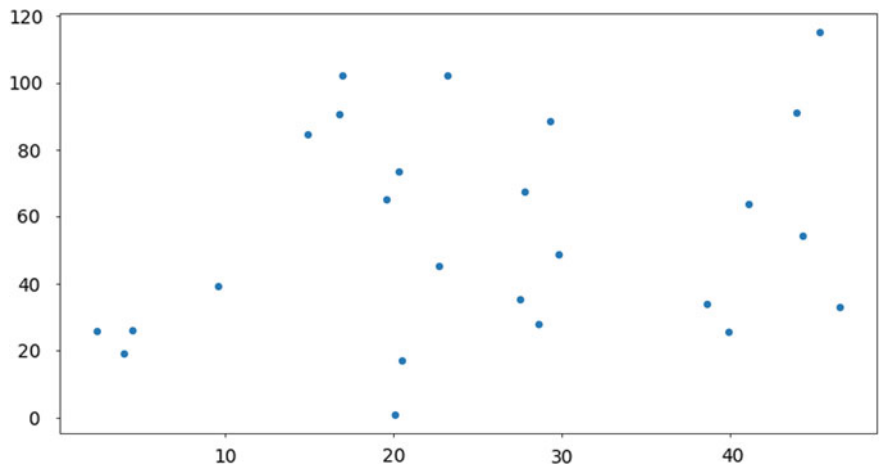
Can a computer model efficiently predict a suitable climbing route? Possibility of finding optimum climbing route from a climbing wall would not be very meaningful until some fundamental data about climbers physique is considered along with shortest path finding algorithms available. The overall process consists of hold generation, compute stride length and adjacent holds, finding the various routes between start and stop points and to find the shortest climbing route as well as alternative climbing route and the difficulty level (Fig. 1).

### 3.1 Random Hold Generation

In traditional natural rock, climbing holds can be a projected section of rock or a small crack which got created naturally in the wall. In natural rock, the hold's



**Fig. 1** Climbing route determination process flow



**Fig. 2** Random hold generation

position in a rock section is random and cannot be compared exactly with another rock section. However, climbing route in an artificial wall is much more predictable when compared to a natural wall. In Moon-board or Grasshopper training board, the holds are placed in matrix fashion and climber includes only the predefined holds given by fellow climbers or route setters. In this implementation, 25 climbing holds being generated randomly on an artificial vertical wall of size 48 inch X 120 inch. The lower most hold is considered as the climbing start point and topmost being the final climbing point (Fig. 2).

**3.2 Determine Stride Length**

A climber can stretch his/her hand and legs as per the stride length to shift the body from one hold to adjacent hold. In best case, climber finds adjacent holds within the stride length. But if there is no holds found then climber needs to stretch legs to a holds more than stride length but if a hold is too far, the climber will not be able to reach and need to take alternate route. Hence, stride length plays a significant role in computing adjacent holds in case of the dynamic move from one hold to another. Generally, as per the physical standard of an average person, the stride length is approximately 0.42 of the person’s height (Table 1) [5].

**Table 1** Height to STRIDE length calculation

Climber's height (In)	Stride length (In)
45	18.9
60	25.2
65	27.3
70	29.4

**3.3 Determine Adjacent Holds**

The characteristics of a climbing wall and the holds position should satisfy the following conditions:

- There should be various routes as per the height of climbers. The route which is suitable for 40 inch tall climbers (Toddlers) should be different from the route designed for 65 inch tall climbers.
- There should be multiple alternate routes for climbers to give an essence of different challenge.
- The route should avoid using closely positioned holds.

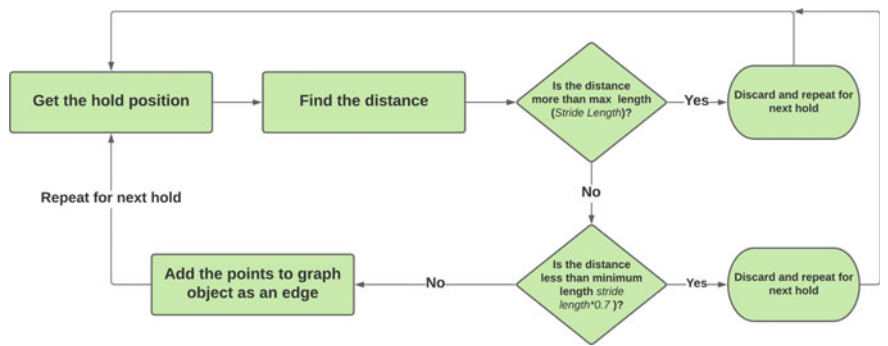
We computed all the hold's distance from each hold irrespective of its position. We considered a climber maximum can stretch his/her leg is the stride length. That means during climbing, if the stride length being 29.4 inch (for a height of 70 inch climber), the maximum stretch of leg would be 29.4 inch. Any hold which is placed beyond stride length would be out of climber's reach and that hold would not be considered for the next move. We discarded the holds which being placed too closely. To achieve that we enforced a rule that a climber has to shift leg by 70% of stride length. That means a climber has to choose a hold no closer to 20.58 inch if the stride length being 29.4 inch. Hence, an adjacent hold has to be within 20.58 inch and 29.4 inch for a height of climber being 70 inch. 70% distance being considered optimum to avoid computational overhead by restricting more number of similar holds and possibility of generating multiple similar climbing routes.

Minimum distance between holds = Stride length X 0.7  
Maximum distance between holds = Stride length

The output list contains the details of the entire nodes and all the possible adjacent holds and the corresponding distance (Fig. 3; Table 2).

**3.4 Finding Different Possible Routes**

The distances from one hold to another can be considered as the vertices of an undirected weighted graph where the holds itself being the edge. The length of the edge is the weight of the edge. To generate the routes, we have applied Dijkstra shortest



**Fig. 3** Adjacent Hold computation process

**Table 2** Adjacent Hold Data

pt_from_X	pt_from_Y	pt_to_X	pt_to_y	pt_from	pt_from	pt_to	Distance(in)
27.5	35.4	4	19.1	pt0	pt20	pt20-pt0	28.59965035
9.6	39.2	4	19.1	pt14	pt20	pt20-pt14	20.8655218
28.6	28	4	19.1	pt22	pt20	pt20-pt22	26.16046636
20.1	0.9	4	19.1	pt23	pt20	pt20-pt23	24.29917694
4	19.1	28.6	28	pt20	pt22	pt22-pt20	26.16046636
20.1	0.9	28.6	28	pt23	pt22	pt22-pt23	28.40176051
2.4	25.9	28.6	28	pt24	pt22	pt22-pt24	26.28402557

path algorithm on the results obtained in Sect. 3.2. The algorithm is implemented in Python using NetworkX package. So, we have a definite set of vertices and edges and the program will find.

- The shortest path between start and stop points using Dijkstra algorithm.
- The alternate paths (set of edges) between start and stop points.

## 4 Experimental Results

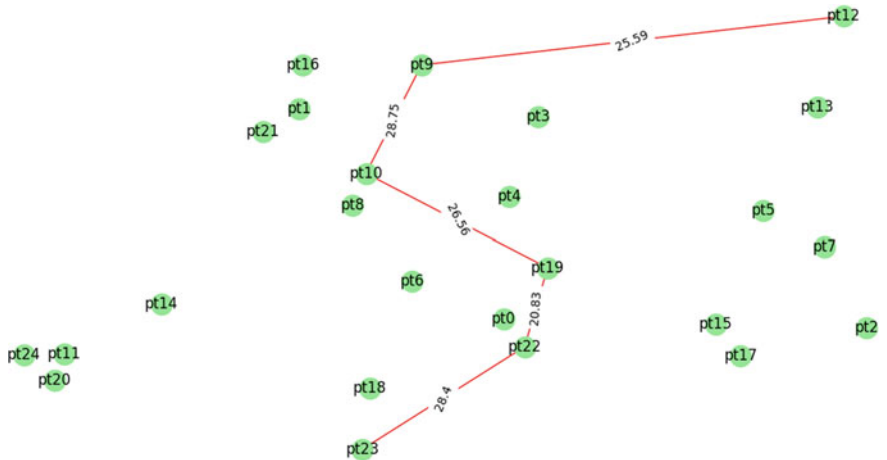
### 4.1 Results

The model is simulated for climbers of height 60 to 72 inch. We observed significant change in the route with the change in height. As the climbers height gets increased the stride also being increased and the program excludes more holds and choose the route having minimum distance. A small change in the climber’s height is computing entirely different route, which is making an artificial climbing wall a chaotic one (Table 3).



**Table 3** Shortest path for different climbers

Climber Height	Climbing Route	Distance Covered	No. of Hops
60	pt23--> pt20--> pt14--> pt19- > pt4- > pt21- > pt9- > pt13- > pt12	174.5781713	8
65	pt23--> pt20--> pt14--> pt19--> pt8--> pt21--> pt9--> pt12	151.8650114	7
70	pt23--> pt22--> pt19-> pt10-> pt9--> pt12	130.1297466	5
72	pt23--> pt11-> pt6--> pt4--> pt13--> pt12	131.4739475	5



**Fig. 4** Shortest climbing route

The algorithm provides shortest path as well as possible alternate routes. Below are the examples of the shortest climbing path for a climber of height 70 inch (Fig. 4).

The below table shows the possible climbing routes for a climber of height 70 inch (Table 4).

The image shows unique alternate paths generated by the algorithm. The paths with same number of hop counts are different and traversed length also not same (Fig. 5).

### 4.2 Difficulty Level

Rock climbing or artificial wall climbing is a challenging sport and route setters have to consider their experience to set up a route and provide gradation to the route. As per UIAA (International Climbing and Mountaineering federation), grade is a kind of estimation of the level of “mental and physical stress” required to pass the route.

Table 4 Possible climbing routes

Route	Path	Distance (in)	Hops
1	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt10->pt9->pt12	330.8993342	13
2	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt10->pt9->pt13->pt12	352.6920553	14
3	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt10->pt16->pt13->pt9->pt12	383.3087742	15
4	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt10->pt16->pt13->pt12	358.2190892	14
5	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt13->pt9->pt12	330.0865002	13
6	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt13->pt12	304.9968153	12
7	pt23->pt20->pt0->pt7->pt2->pt6->pt4->pt1->pt8->pt3->pt5->pt16->pt10->pt9->pt12	393.1068381	15

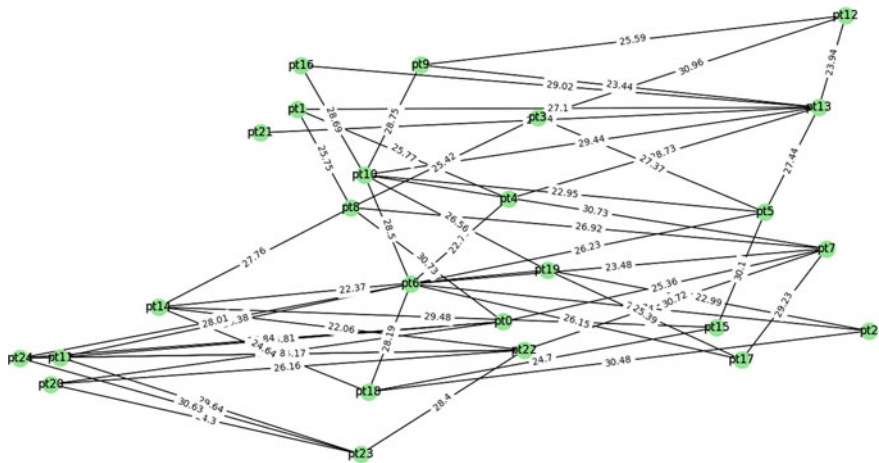


Fig. 5 Possible climbing routes

In this implementation, the following parameters being considered to find difficulty level of climbing route.

- Deviation of the route from the shortest route.
- Increase in Number of hops, i.e., climber has to do more moves.
- Increase in route distance while numbers of holds are less. This means the climber do not find holds nearby to his stride length and has to stretch legs to the farthest possible distance. More the leg stretch, more strenuous the route becomes.

The route difficulties can be explained by following variations.

1. Route 1 has 330 inch distance covered with 13 hops and route 2 has 352 inch covered with 14 hops. In this scenario, route 1 is better route as it takes less hop and less distance.
2. Route 2 has 352 inch distance covered with 14 hops, and route 4 has 358 inch covered with 14 hops as well. In this scenario, route 2 being better route than route 4.

Difficulty Level	Hops	
Distance	Easy	Moderate
	Moderate	Difficult

### 4.3 Analysis and Discussion

This paper explores the possible climbing routes and its difficulty level by the number of hops and the distance between the holds. Climbing route setting can be widely used in climbing gyms or practice walls to enable numerous routes for climbers for all age and height. This paper emphasizes mostly on the route finding rather than difficulty level. Rock climbing is a challenging sport which requires climber's body part movement, stress on hand and feet to balance body. The size of the holds, angle of the holds from the climber, slope of the wall also affects the difficulty level of the route. These elements need to be considered as future study of computer modeling a climbing route. In a classical 3 point climbing, a climber supports his/her body to the wall by balancing two legs and one hand or by two hands and one leg. The movement of arm needs to be considered for future work for difficulty calculation. The possibility of collecting data related to stress on hand and foot can be thought of.

## 5 Conclusion

This paper describes the complete understanding of route setting process of a climbing wall where holds are placed in random fashion. This implementation focusses on an unique approach of route finding according to climbers height and his/her ability to shift to adjacent holds. The graph theory based approach is unique in this area, and the simulation yields convincing results. The work further could be extended by adding different types of holds with different angle and to analyze the difficulty level of the route.

**Acknowledgements** We would like to thank Dr. Nilankur Dutta of Institute Lumiere Mateire, Universite Claude Bernard Lyon-1, France, for his valuable review of the paper and feedback.

## References

1. Stapel F (2020) A heuristic approach to indoor rock climbing route generation. University of Twente
2. Salewa cube: highlight for the Bolzano climbing scene. Salewa Cube website, May 2011
3. Kempen L (2018) A fair grade: assessing difficulty of climbing routes through machine learning. Formal methods and tools, University of Twente
4. Naderi K, Takatalo J, Lipsanen J, Hamalainen P (2018) Computer-aided imagery in sport and exercise: a case study of indoor wall climbing. Toronto
5. <https://www.verywellfit.com/set-pedometer-better-accuracy-3432895>
6. Johnson DB (1973) A note on Dijkstra's shortest path algorithm. J ACM 20(3)
7. Chen JC (2003) Dijkstra's Shortest Path Algorithm, J Formalized Math 15

8. Mandelli G, Angriman A, Scales of difficulty in climbing. Central School of Mountaineering—CAI, Italy ; (translation by G. Coppadoro— CAI)
9. [https://www.theuiaa.org/documents/sport/THE-SCALES-OF-DIFFICULTY-IN-CLIMBING\\_p1b.pdf](https://www.theuiaa.org/documents/sport/THE-SCALES-OF-DIFFICULTY-IN-CLIMBING_p1b.pdf)

# Integration of Attribute-Based Access Control in Microservices Architecture



Amandeep Singh, Vinay Raj, and Sadam Ravichandra

**Abstract** Microservice is a software architecture with significant improvements over service-oriented architecture (SOA). The definition of microservices states that each service should perform only a single task, and hence, applications will have more services. As the number of services increases, different issues arise w.r.t testing, monitoring, governance, and security. When services communicate in microservices architecture, new challenges arise in areas like containers (deployment environment), data, permissions (authorization and authentication), and network. We will focus on permission issues related to authentication and authorization. Role-based access control (RBAC) model was used in microservices to ensure authorization. However, there are some security issues in RBAC, i.e., role explosion and segregation of duty. These issues also have an effect on the security policies of microservices. We aim to propose a solution to solve these problems by providing attribute-based access control for microservices architecture. We evaluate the performance of the proposed model by measuring metrics such as response time, throughput, and data transferred. From testing, it is proved that the system is reliable.

**Keywords** Microservices · Service-oriented architecture · Role-based access control · Attribute-based access control

## 1 Introduction

Earlier, the main focus of many implementations is to ensure authentication. The authorization has always been overlooked. But, nowadays both, authentication and authorization are given equal importance, and most implementations now consider

---

A. Singh (✉)  
National Institute of Technology, Warangal, India

V. Raj  
BVRIT HYDERABAD College of Engineering for Women, Hyderabad, India

S. Ravichandra  
National Institute of Technology, Warangal, India

access control module as their integral part [5]. Role-based access control (RBAC) [14] became the standard for access control. This model uses the principle of roles. This allows organizing access control based on roles and permissions mapped with those roles. RBAC suffers from many problems with the increase of access control use cases. The main problem was “Role Explosion” [17]. This problem arises when the number of roles and permissions associated with those roles increases to such a point when the system becomes impossible to manage. The other problem RBAC faces is “Segregation of Duty.” RBAC does not have any mechanism to solve this problem and therefore results in altering the requirements of the system. Attribute-based access control (ABAC) solves the above-mentioned problems. With the use of attributes and policies, ABAC can implement relationships which makes it more flexible and provide thorough access. Relationships help to solve the problems of “Role Explosion” and “Segregation of Duty.” ABAC makes the process of authorization easy by using the user’s attributes directly, and there is no need to worry about adding and deleting permissions associated with the user [16].

In this paper, we propose a microservices architecture with attribute-based access control as its authorization module. We measure the performance of the proposal by measuring metrics such as response time, throughput, and data transferred. Siege tool is used to record the values of performance metrics. The result shows that microservices architecture with ABAC performs better than microservices architecture with RBAC.

## 2 Attribute-Based Access Control (ABAC)

This section describes the ABAC model, i.e., the attributes, the policies, and how these policies are applied in the authorization module.

### 2.1 Attributes

The ABAC model allows authorization to be defined in the form of policies, and these policies consist of attributes. There are three types of attributes used for implementing access control [16], which are as follows:

- Subject Attributes: A subject is defined as an entity which acts on an object.
- Resource Attributes: A resource is an object on which subject acts.
- Environment Attributes: Most access management measures have so far largely overlooked environmental attributes.

ABAC completely incorporates the functionality of the RBAC model by treating role as a characteristic of a principal. Moreover, it can easily manage MAC requirements by being able to understand resource attributes such as sensitivity detail. As a consequence, we believe ABAC is a product of existing models of access control,

and it outperforms their functionality. The representation of policy is more semantically complete and expressive, and it can be more thorough within ABAC since any combination of attributes can be used.

## 2.2 ABAC Policy Formulation

In this subsection policy model of ABAC is defined and explained with examples.

- Subjects, resources, and environments are depicted through S, R, and E, respectively;
- $SAttr_k (1 \leq p \leq P)$ ,  $RAttr_m (1 \leq q \leq Q)$ , and  $EAttr_n (1 \leq r \leq R)$  are the designated attributes to subjects, resources, and environments, respectively;
- The attribute assignment relations for subject s, resource r, and environment e, are  $Attr(s)$ ,  $Attr(r)$ , and  $Attr(e)$ , respectively:

$$Attr(s) \subseteq SAttr_1 \times SAttr_2 \times \dots \times SAttr_P \quad (1)$$

$$Attr(r) \subseteq RAttr_1 \times RAttr_2 \times \dots \times RAttr_Q \quad (2)$$

$$Attr(e) \subseteq EAttr_1 \times EAttr_2 \times \dots \times EAttr_R \quad (3)$$

For assignment of value to attributes, we have used function notation. For example,

$$Role(s) = \text{“Service Consumer”} \quad (4)$$

$$Owner(r) = \text{“XYZ”} \quad (5)$$

$$Date(e) = \text{“21 – 05 – 2021”} \quad (6)$$

- Policy rule is defined as a boolean function of attributes (subject s, resource r, and environment e).

$$\text{Rule : Access}(s, r, e) \leftarrow (Attr(s), Attr(r), Attr(e)) \quad (7)$$

If the boolean function results true, then the resource r can be accessed by subject s; otherwise, not.

- A policy rule base, also known as a Policy Store, may contain a large number of policy rules that cover a wide range of subjects and resources within a security domain.

Now we'll look at how attributes can be used to create various policy rules. The basic rules are those that consist of attributes (subject, resource, and environment) that are not correlated to each another and these are widely used types of policy rules. In this case, policy rules normally assign constant values to attributes. A rule such as



“Users with the position of ‘Manager’ only can access the ‘Purchase’ web service” could be written as:

$$R_1 : \text{access}(s, r, e) \leftarrow (\text{Role}(s) = \text{'Manager'}) \wedge (\text{Name}(r) = \text{'Purchase'}) \quad (8)$$

The above rule also demonstrates that the ABAC model can easily accommodate traditional RBAC rules. The ABAC model, on the other hand, can model far richer access control semantics. For example, we can provide the following rule to ensure that a resource can only be accessed by its owners:

$$R_2 : \text{access}(s, r, e) \leftarrow (\text{UID}(s) = \text{Resource Owner}(r)) \quad (9)$$

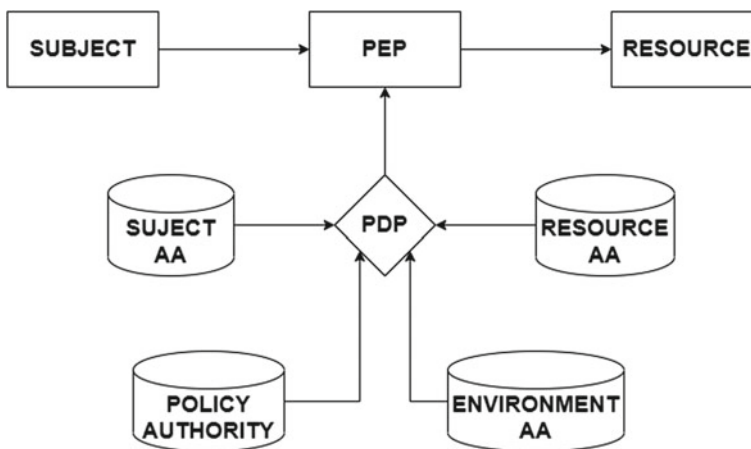
The evaluation of policy rules can be reduced to the evaluation of first-order logic expressions given attribute assignments, and forward chaining can be easily implemented.

### 2.3 Authorization Architecture

This subsection explains the components of attribute-based authorization architecture. The components are shown in Fig. 1.

The diagram shows the components of ABAC:

- The Attribute Authorities (AA) are in charge of developing and maintaining attributes.
- The Policy Enforcement Point (PEP) is in charge of requesting and implementing decisions corresponding to the input query.



**Fig. 1** Components of authorization system

- The Policy Decision Point (PDP) is in charge of assessing the relevant policies and making the authorization decision (permit or deny).
- The Policy Authority Point (PAP) is in charge of creating and managing access control policies.

### 3 Related Work

In recent years, a lot of effort has gone into dealing with security issues in the microservices architecture [18]. Containers, data, permissions, and the network are the four aspects of security in microservices architecture [6]. We focused on authentication and authorization permission issues. Compromise of a single service can affect other services and eventually leading to the failure of the whole system. Therefore, the authenticity of each service should be ensured by microservices architecture. When a service receives a message, it must also determine whether the message is valid and whether the source service has valid authority. ABAC in SOA improves the scalability and alterability [1]. Ni Dan et al. used the fact that ABAC can use security attributes to define policies that enable fine-grained access control. Their system used security assertion markup language (SAML) to pass security tokens. Identity-based access control (IBAC) cannot be used in SOA because it cannot withstand the scalability, security, and manageability of SOA [2]. Alan H. Karp presented that IBAC results in a tightly coupled system, and therefore, achieving authorization is difficult, whereas RBAC compliments the SOA and helps to achieve authorization. Sabbari et al. [3] addressed the issue that if the number of services increases and their communication gets complicated, then SOA faces security threats. To solve this problem, they presented an architecture that controls the resources based on the ABAC model and uses two standard languages XACML and SAML.

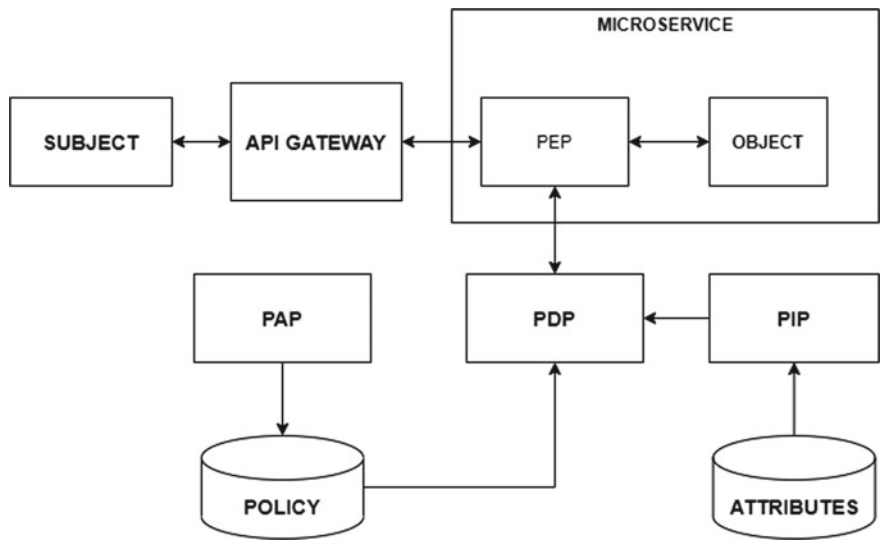
Yuan et al. [4] implemented ABAC on SOA with the help of open-source tools for ensuring web services. They explained the access control challenges faced by SOA and how ABAC solves them. They also addressed the differences between ABAC and RBAC, which clearly showed the advantages of ABAC. Pereira-Vale et al. [5] described the mapping to identify the security mechanisms for microservices architecture. Through their mapping, it is concluded that most of the security mechanisms were focused on authentication and authorization. All mechanisms dealt with stopping the security attacks, but none of them focused on how to recover from a security attack. Yu et al. [6] presented work related to security threats for microservices-based fog applications and stated that security issues arise in four areas like containers (deployment environment), data, permissions (authorization and authentication) [9], and network security. They proposed a solution to solve the security issues in services communication. Yang et al. [7] designed an access control system for microservices architecture based on OAuth 2.0 protocol. They used JSON Web Token (JWT) to ensure authorization and a gateway to ensure authentication. Triartono et al. [8] proposed an architecture implementing role-based access control on OAuth 2.0 using the Laravel framework. They proved that their system is reliable by using siege

testing tools. Tran Thanh et al. [13] described a framework that enabled the consideration of security and privacy across service’s lifecycle by taking into account new technologies and new architectural design (i.e., microservices architecture).

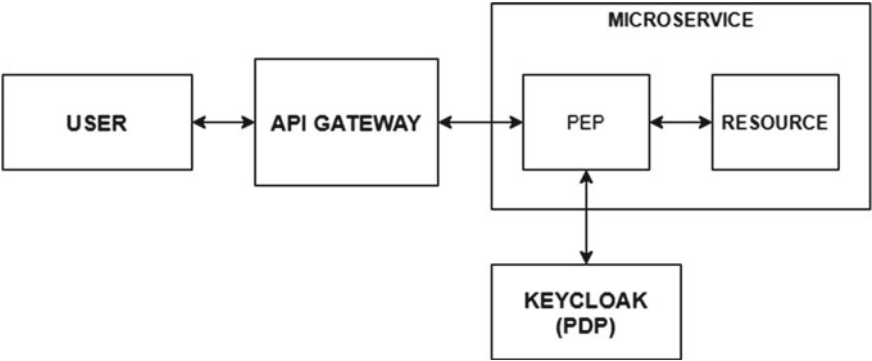
### 4 Architecture

This section describes the proposed microservices architecture with attribute-based access control for authorization among services. The proposed architecture is shown in Fig. 2.

- The working of the system is as follows:
- 1. Policy Administration Point (PAP) is used to describe access control rules (policies).
  - 2. These rules, along with the associated attributes are sent to the Policy Decision Point (PDP).
  - 3. When the subject connects with microservice via API gateway, the code at microservice connects with centralized Policy Decision Point (PDP).
  - 4. PDP produces a decision by comparing the input with rules and attributes. This decision is then sent back to the PEP.
  - 5. Microservice implements the decision with the help of PEP.



**Fig. 2** Microservices architecture with ABAC



**Fig. 3** Experimental setup

## 5 Implementation

We implemented a proof of concept by using Keycloak as an authorization server. We have conducted this experiment on a machine with Ubuntu 20.04 and 8 GB of RAM. Keycloak supports fine-grained access control policies and can be combined easily with ABAC. For authentication, we have used an API gateway that uses the session information from the user’s end and authenticates accordingly. The policies are written in XACML [15]. The experimental setup is shown in Fig. 3.

## 6 Results

This section will discuss the results obtained while testing the system using the siege tool. The metrics that will be measured are response time, throughput, and data transferred. The values of the metrics are observed by varying the number of users, i.e., 50, 100, 150, 200, and 250. The response time measurement is shown in Fig. 4. This shows that as the number of users increases, response time also increases which signifies that the system is stable. The throughput measurement is shown in Fig. 5. This shows that system is stable and reliable. The data transferred measurement is shown in Fig. 6. In this, it is observed that the system performance is slowing down when the number of users reached 250 which is caused by loss of connections.

The system performed quite well and proved to be reliable. Although the system incurred a loss of data as it reached 250 users, we consider that it can be resolved as we scale the resource.

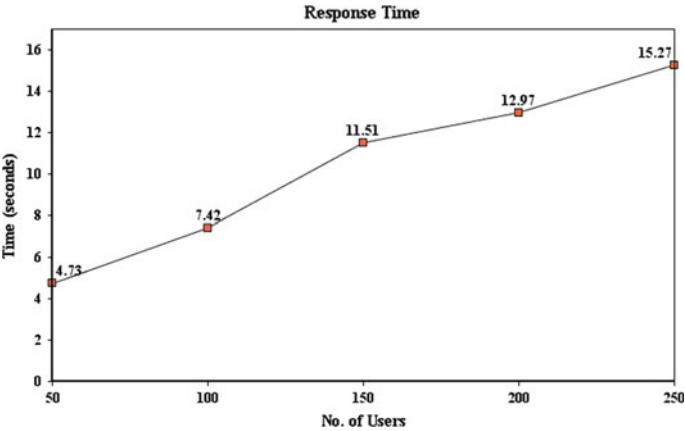


Fig. 4 Response time measurement in proposed architecture

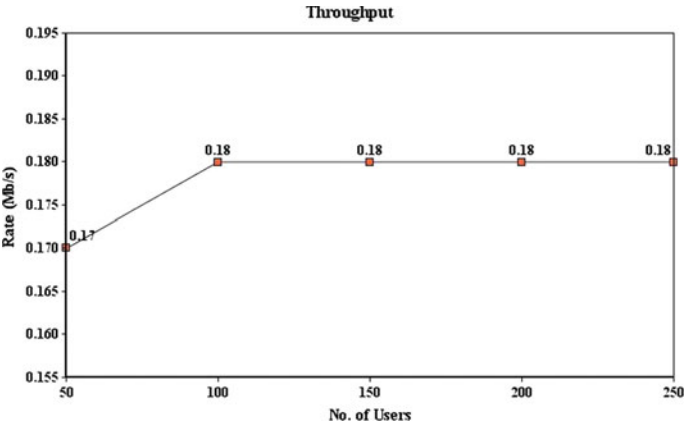


Fig. 5 Throughput measurement in proposed architecture

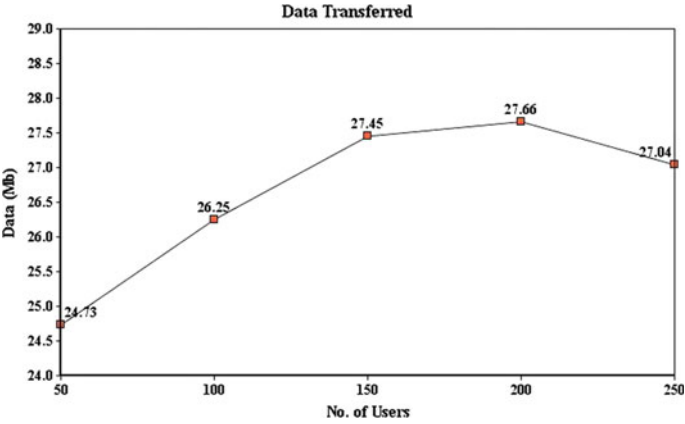


Fig. 6 Data transferred measurement in proposed architecture

## 7 Conclusion

In this paper, we highlighted the problems of RBAC (role explosion and segregation of duty). To solve these problems, we proposed a microservices architecture with attribute-based access control as its authorization module. We have used siege tool to measure performance metrics (response time, throughput, and data transferred). By comparing the performance of ABAC in microservices with RBAC in microservices [8], we can conclude that microservices architecture with ABAC performs better than microservices architecture with RBAC. Our solution relies on trusting the server, but an attack on a server can compromise the entire system.

## References

1. Dan N, Hua-Ji S, Yuan C, Jia-Hu G (2012) Attribute based access control (ABAC)-based cross-domain access control in service-oriented architecture (SOA). In: 2012 international conference on computer science and service system, Nanjing, 2012, pp 1405–1408. <https://doi.org/10.1109/CSSS.2012.354>
2. Karp AH (2006) Authorization-based access control for the services oriented architecture. In: 4th international conference on creating, connecting and collaborating through computing (C5'06), Berkeley, CA, , pp 160–167. <https://doi.org/10.1109/C5.2006.9>
3. Sabbari M, Alipour HS (2011) Improving attribute based access control model for web services. In: World congress on information and communication technologies. Mumbai, pp 1223–1228. <https://doi.org/10.1109/WICT.2011.6141423>
4. Yuan E, Tong J (2005) Attributed based access control (ABAC) for web services. In: IEEE international conference on web services (ICWS'05), Orlando, FL, pp 569. <https://doi.org/10.1109/ICWS.2005.25>
5. Pereira-Vale A, Márquez G, Astudillo H, Fernandez EB (2019) Security mechanisms used in microservices-based systems: a systematic mapping. In: 45 Latin American computing conference (CLEI). Panama, Panama, pp 01–10. <https://doi.org/10.1109/CLEI47609.2019.235060>
6. Yu D, Jin Y, Zhang Y, Zheng X (2019) Nov 25 a survey on security issues in services communication of Microservices-enabled fog applications. *Concurr Comput Pract Exp* 31(22):e4436
7. ShuLin Y, JiePing H (2020) Research on unified authentication and authorization in microservice architecture. In: 2020 IEEE 20th international conference on communication technology (ICCT), pp 1169–1173. <https://doi.org/10.1109/ICCT50939.2020.9295931>
8. Triartono Z, Negara RM, Sussi (2019) Implementation of role-based access control on OAuth 2.0 as authentication and authorization system. In: 2019 6th international conference on electrical engineering, computer science and informatics (EECSI), Bandung, Indonesia, pp 259–263. <https://doi.org/10.23919/EECSI48112.2019.8977061>
9. Chandramouli R (2019) Security strategies for microservices-based application systems. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP), pp 800–204. <https://doi.org/10.6028/NIST.SP.800-204>
10. Chandramouli R, Butcher Z (2020) Building secure microservices-based applications using servicemesh architecture. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-204A. <https://doi.org/10.6028/NIST.SP.800-204A>
11. Liu B, Yang Y, Zhou Z (2018) Research on hybrid access control strategy for smart campus platform. In: IEEE 3rd advanced information technology, electronic and automation control conference (IAEAC). Chongqing 2018, pp 342–346. <https://doi.org/10.1109/IAEAC.2018.8577828>

12. Ferraiolo D, Chandramouli R, Kuhn R, Hu V (2016) Extensible access control markup language (xacml) and next generation access control (NGAC). In: Proceedings of the 2016 ACM international workshop on attribute based access control (ABAC'16). Association for Computing Machinery, New York, NY, USA, pp 13–24. <https://doi.org/10.1145/2875491.2875496>
13. Thanh TQ, Covaci S, Magedanz T, Gouvas P, Zafeiropoulos A (2016) Embedding security and privacy into the development and operation of cloud applications and services. In: 17th international telecommunications network strategy and planning symposium (Networks). Montreal, QC 2016, pp 31–36. <https://doi.org/10.1109/NETWKS.2016.7751149>
14. Sandhu RS, Coyne EJ, Feinstein HL, Youman CE (1996) Role-based access control models. *Computer* 29(2):38–47. <https://doi.org/10.1109/2.485845>
15. OASIS, The XML access control markup language (XACML) OASIS TC Homepage. [http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=xacml](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml)
16. Hu VC, Kuhn DR, Ferraiolo DF, Voas J (2015) Attribute-based access control. *Computer* 48(2):85–88. <https://doi.org/10.1109/MC.2015.33>
17. Rajpoot QM, Jensen CD, Krishnan R (2015) Integrating attributes into role-based access control. In: IFIP annual conference on data and applications security and privacy. Springer, Cham
18. Raj V, Ravichandra S (2018) Microservices: a perfect SOA based solution for enterprise applications compared to web services. In: 2018 3rd IEEE international conference on recent trends in electronics, information and communication technology (RTEICT) 2018 May 18. IEEE, pp 1531–1536

# Analysis and Forecasting of Cybercrime Incident in India



Kousik Barik, Karabi Konar, Archita Banerjee, Saptarshi Das,  
and A. Abirami

**Abstract** India is gradually proceeding toward being a digital economy. This has led to an increase in changelessness toward cybersecurity. Conforming to the National Crime Records Bureau (NCRB), cybercrime has increased manifolds in the past 20 years. In this work, we have used time series analysis to forecast future trends of cyberattacks. The various motives of cybercrimes and categories are analyzed on statewise basis. Experimental evaluation shows that our model is capable of predicting the future trends of cybercrime incidents. We also graphically represent the suggested future prediction and visualization for the same. Such forecasting is vital for raising awareness among and mitigating future attacks.

**Keywords** Cybercrime · Cybersecurity · Statistical analysis · Motives of cybercrime · Cyberlaws

## 1 Introduction

Over the last few years, India has made significant advancements in building up a digital framework for the digital economy. In 2002, India had 16,765,756 Internet users. As of 2019, India possesses 636.73 million Internet users [1]. The advancement of communication technologies is attributed to an exponential increase in the number of Internet users. The total number of active Internet users is projected to be 974.84 million by 2025 [2]. This growth in user base has led to a boom in the number of cybercrime incidents. Between 2015 and 2019, over 33,000 cybercrimes have been recorded in India. In 2017, India ranked third [3] in comparison with the highest number of cyberthreats detected among the list of countries. India recorded 21,796 [4] cybercrimes in 2017, which is an increase of 77% from 2016 [4].

---

K. Barik (✉) · K. Konar · A. Banerjee · S. Das  
JIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

S. Das  
e-mail: [saptarshi@jisiasr.org](mailto:saptarshi@jisiasr.org)

A. Abirami  
Bannari Amman Institute of Technology, Erode, India



Crime is an undesirable act that results in imposing a legal reprimand for the culprit. Cybercrime primarily entails unauthorized access to systems, stealing of data, and theft of intellectual property. Additionally, child pornography, child sexual abuse, cyberbullying, cyberstalking, online job fraud, online sextortion, vishing, sim swap scam, phishing, spamming, denial of services, and online drug trafficking [2] are of few examples of cybercrimes. Investigating a cybercrime is complex, and most of the countries worldwide are adopting best practices to deal with it. Every country has its own cybersecurity law or Internet law to safeguard all stakeholders. Information Technology Act 2000 [5], framed by the Government of India, deals with cybercrime and electronic commerce in India.

“History often repeats itself” Whatever happens in the history, they are expected to take place also in the forthcoming. Time series data is time-dependent. In the graphical representation, the x-axis illustrates the independent variable and the y-axis illustrates the dependent variable. With time-series data, the independent variable is always time. Autoregressive integrated moving average (ARIMA) [6] is a way of designing time-series data for forecasting based on past values. ARIMA can pick up complex relations by choosing error terms and observations of lagged terms.

In this context, the motivation of this work deals with cybercrime analysis, crime patterns, and generating user awareness so that all end-users understand and report cyber incidents across the country. This paper gives a comprehensive statistical analysis of the cybercrimes registered and types of cyber crimes during 2017–2019 [7]. The remaining paper is formulated as follows; Sect. 1 discusses literature survey. In Sect. 2, we elucidated the statistical forecasting of cyberattacks in India and future prediction of attacks. Section 3 comprises the cyber crimes registered under IPC and IT Act, statewise crime pattern analysis. In Sect. 4, we explored the various motives of the attacker and the reasons behind not reporting cyber incidents. Next in Sect. 5, we discussed the initiative and user awareness program undertaken by the government. Finally, we conclude the paper in 6 and provide various direction for future research.

## 2 Literature Survey

Sreehari et al. [8] raised consciousness among college students at Kochi related to cybercrimes that introduced many safeguards accepted by students while accessing the Internet. Chatterjee et al. [9] showed a conceptual model to determine elements limiting cybercrimes, affirms that “awareness of cybercrimes” affects the current usage of technology to respond cybercrimes in smart cities in India. Gunjan et al. [10] presented a review study based on cybercrime in India in 2013, which represented different views of cybercrime with case studies and concluded cybercrime is more terrible and destructive in connection with traditional crimes. Singh and Rastogi [11] extended different levels of cybercrimes, build user perception about the usage of the Internet accurately, makes user understanding, has precautions from turning into an easy target, knowledge cyberrules to be extended broadly

so that victims can choose actions accordingly. Kshetri [12] developed a structure that defined the relationship between formal and informal institutions, significant to cybersecurity and cybercrime in growing countries. Kumar [13] suggested a study on aging cybercrimes in India in 2016, showed the rise across the country, and leadership started by the Government of India to encounter cybercrimes. Datta et al. [14] presented a comprehensive study on cybercrimes in India, fraud cases are rising, the sufferers are usually in the age group of 20 to 29, particularly children and women are becoming affected, proposed a user awareness program to counter cybercrimes. Various government initiatives also exist in this regard. The Government of India, Ministry of Law established Information Technology Act, 2000 (Indian Parliament: No 21 of 2000) notified on 17 October 2000 [15], permitting statutory approval to all transactions done through electronic data exchange, related to cybercrime is a punishable offense, the bill enhanced in 2008 as Information Technology (Amendment) Act, 2008 with further corrections and additions. Dataset produced by the NCRB contains cybercrimes under offenses registered under Information Technology Act, 2000 [15] and Offenses registered under the Indian Penal Code (IPC) [7].

### 3 Statistical Forecasting of Cyberattacks in India

The present work aims to predict future cybercrime attacks in India by analyzing past cybercrime data to predict future values. In this work, we have used cybercrime data from 2002 to 2019 to predict the future trend of cyberattacks, i.e., for 2020 to 2025, and to compare the predicted values with the actual ones of the attack for the current years by using autoregressive integrated moving average (ARIMA) [6] model.

We use the ARIMA model on yearly cybercrime offenses between 2002 and 2019 in India. First, we checked the stationarity of the time series first before implementing ARIMA. Here, the time series is non-stationary, so we convert it into a stationary one, where mean, variance, and covariance are constant with time. The stationarity of a series is verified using augmented Dickey–Fuller tests [16].

Next, we examined the correlogram and partial correlogram [17] with the use of “acf()” and “pacf()” in R. We need to decompose the seasonal and non-seasonal data to remove the noise and seasonal trend from the time series. For constructing the appropriate ARIMA model, we used the auto-arima package in R Studio that finds the best fit.

$$\text{ARIMA}(2, 0, 1) y_t = a_1 y_{t-1} + a_2 y_{t-2} + b_1 \epsilon_{t-1}$$

$$\text{ARIMA}(3, 0, 1) y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + b_1 \epsilon_{t-1}$$

$$\text{ARIMA}(1, 1, 0) \Delta y_t = a_1 \Delta y_{t-1} + \epsilon_t$$

$$\text{ARIMA}(2, 1, 0) \Delta y_t = a_1 \Delta y_{t-1} + a_2 \Delta y_{t-2} + \epsilon_t$$

$$\text{where } \Delta y_t = y_t - y_{t-1}$$

ARIMA equations:

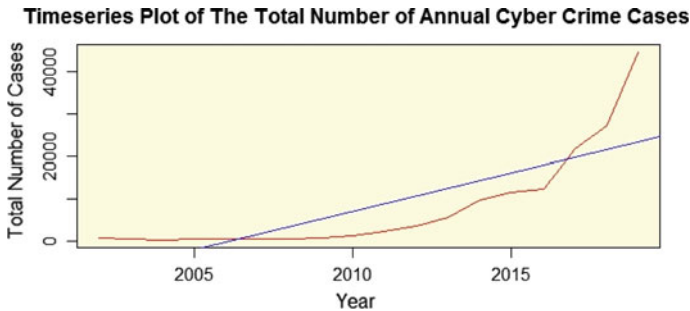
$$\begin{aligned} \text{ARIMA}(1, 1, 0) : y_t &= a_1 \Delta y_{t-1} + \epsilon_t \\ \text{ARIMA}(2, 0, 0) : y_t &= a_1 y_{t-1} + a_2 y_{t-2} + \epsilon_t \\ \text{ARIMA}(2, 1, 1) : \Delta y_t &= a_1 \Delta y_{t-1} + a_2 \Delta y_{t-2} + b_1 \epsilon_{t-1} \\ \text{where } \Delta y_t &= y_t - y_{t-1} \end{aligned}$$

Next, we estimated the ARIMA model by applying the values of  $p$ ,  $d$ ,  $q$  and determine the appropriateness, to check the nature of the residuals to see in case of white noise. The finest model with a well-behaved residual is accepted. The final step is the forecasting of the model with the desired number of upcoming predictions. There are various existing statistical forecasting methods, e.g., ARMA. If we use ARMA model, then we need to determine  $p$  and  $q$  using acf or pcf, and then after passing ARMA( $p$ ,  $q$ ) the model will be created. But with ARIMA, we just have to pass the data into it and then ARIMA will by itself estimate the most optimized  $p$  and  $q$  and then on that basis model will automatically be created. ARIMA is very flexible compared to other models.

We have transformed the dataset into a time-series format to visualize distinctly presented in Fig. 1, fitted a line that represents a spike in the number of cybercrime cases over the years. During the initial years, the growth of cybercrime cases was linear, but during recent years, the exponential growth of the cybercrime records in the country.

We established the ARIMA ( $p$ ,  $q$ ,  $d$ ) model [16] where  $p$  describes the number of autoregressive terms,  $d$  illustrates the number of non-seasonal differences required for stationery, and  $q$  describes the number of lagged forecast errors in the prediction equation, which serves to enhance estimate out the data which is needed to bring out future predictions, represented in Fig. 2.

Our model automatically picks up the most optimized ARIMA model, ARIMA (0,2,1), a very lower value of residuals which signifies that the difference between original values and predicted values is very less. We conclude that our model has predicted the future outcomes correctly. The residuals, i.e., the errors in prediction from 2002 to 2019 are 0.002993899, -0.010188486, 0.212104778, 0.708169128, -0.061345687, 0.236063848, -0.275388858, 0.457544657, 0.450083862,



**Fig. 1** Time-series plot of total number of annual cybercrime cases

Model	Predictions	Model	Predictions	Model	Predictions
ARIMA(2,2,2)	23.64101	ARIMA(0,2,0)	12.98403	ARIMA(1,2,0)	12.93384
ARIMA(1,2,1)	15.65083	ARIMA(0,2,2)	15.64819	ARIMA(1,2,2)	19.07164
ARIMA(0,2,1)	12.5977				

Fig. 2 ARIMA model

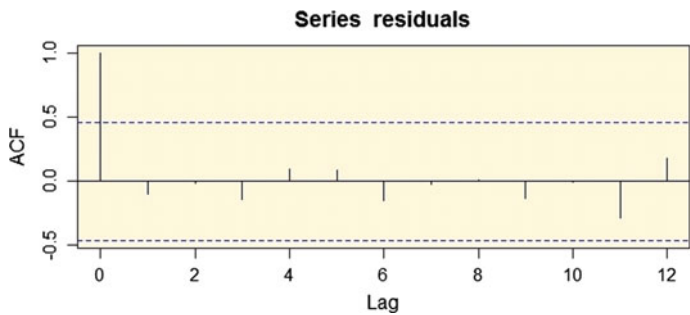


Fig. 3 Error in irregular pattern

0.084171112, −0.024012575, 0.030015115, 0.045786502, −0.317135646, −0.273934218, 0.381951402, −0.168839226, 0.189316070, respectively.

We plotted acf of the residuals. In the case of linear regression, we used to plot the error against the fitted values to check if they are random or not. But as this is time-series data, we need to check whether the errors are stationary. While plotting the errors, we observed the lag increases; going toward being stationery. This shows that errors in terms of regression are independent, randomly scattered, in terms of time series, the errors are in an irregular pattern. We consider this model to be an excellent model, and the fit is good, represented in Fig. 3.

In the Box–Ljung test [18], the null hypothesis is errors are not autocorrelated. Because the p-value is higher than the level of significance, we will not reject the null hypothesis, which is our key requirement that at the end, errors not be auto-correlated. Errors should be random and we have achieved that. Hence, we can conclude that it is a good fit. Looking at the low p-value from the Box–Ljung test, we conclude our model is relatively accurate from the ARIMA model, parameters (0,2,1) adequately fit the data. Box–Ljung test: data: model\$residual, X-squared = 0.20805, df 1, p-value = 0. 6483. Subsequently, we performed forecasting using the forecast function in R. We projected data for the upcoming 6 years along with the confidence intervals, represented in Fig. 4.

Figure 5 represents the forecast plot above. The dark blue color around the line signifies an 80% confidence interval; the gray portion signifies a 95% confidence interval. This means we are 80% confirmed that if our data deviates then it will deviate within the dark blue region, and 95% confirmed that even if the data deviates, it will deviate within the gray region. Hence, our forecast will be in range. Our model predicts the future attacks for the years 2020, 2021, 2022.

Year	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2020	11.10727	10.703311	11.51123	10.489469	11.72507
2021	11.51026	10.771136	12.24939	10.379867	12.64066
2022	11.91325	10.798878	13.02763	10.208964	13.61754
2023	12.31625	10.785814	13.84668	9.975653	14.65684
2024	12.71924	10.734312	14.70416	9.683556	15.75492
2025	13.12223	10.646867	15.59759	9.336489	16.90797

Fig. 4 Projection of cybercrime for upcoming 6 years

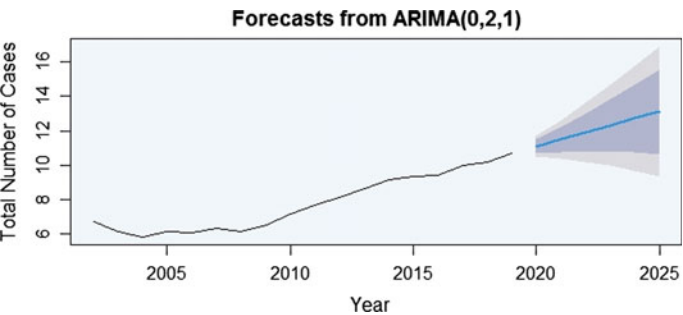


Fig. 5 Forecast from ARIMA (0,2,1)

Time Series: Start = 2020, End = 2025, Frequency = 1, 66,653.95 99,733.96 149,231.41 223,294.18 334,113.92 499,932.93.

We furnish the actual versus predicted values in graphical plotting, shown in Fig. 6. We can envision the actual and predicted values; a comparison study is made subsequently. Our model fits pretty well as the variation between current and forecast

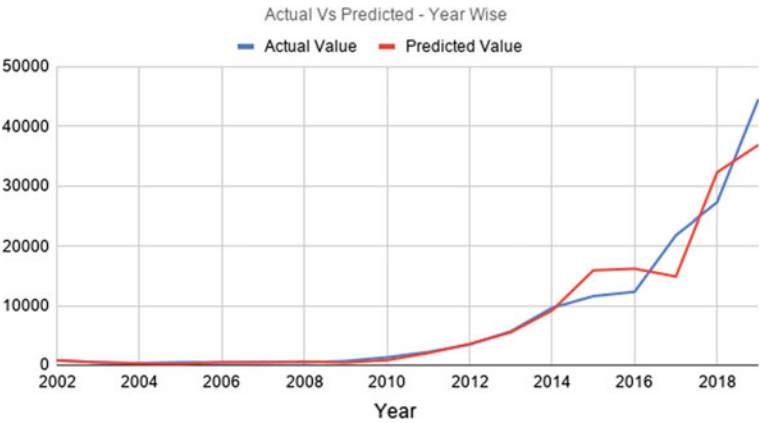


Fig. 6 Actual versus predicted value

values is less. The reason behind this is if we estimate an already given point, then we call it a fit. But if we estimate an unknown time point, then it is called prediction. For fitting or predicting, we do not need dataset; instead, we need a model.

Based on some given data points, that model is prepared. Now the model is prepared as a dependent variable and time points. In that case, if we place any other time, we will get an output, be it any given time point or any future time point. When we place a time point from our given dataset, we are getting a value that we call fit or fitting. Whereas when we put an unknown time point, we are calling it a prediction.

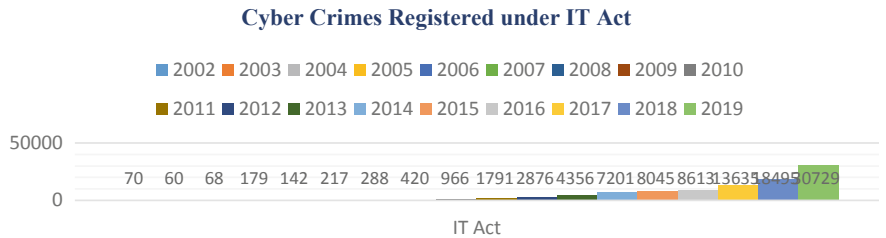
## 4 Cybercrimes Registered

### 4.1 Cybercrimes Registered Under IT Act

Cybercrime is increasing faster in India over the previous few years, prominent from the statistical presentation in Fig. 7 from 2002 to 2019.

Based on NCRB reported data in 2019, 2018, and 2017, the dominant 6 cybercrime heads under IT Act are identity theft (sec 66C), computer-related offenses (sec 66), cheating by personation by utilizing computer resources (Sec.66 + D), other sections of the IT Act [19], publishing or transmitting obscene material in electronic form, publishing or transmitting of material containing sexually explicit act in electronic format (Sec.67A) [19].

Sec 66c of the IT [19] act accords with offenses related to identity theft; registered cases under this head are 3724 in 2017, 6688 in 2018, 12,255 in 2019. Sec 65 of IT [19] acts deals with computer-related offenses related to crime heads, lose damage to computer systems, offenses because of ransomware, offenses other than ransomware, registered cases is 3508 in 2017, 3969 in 2018, 4467 in 2019, divided into two parts: offenses with ransomware and offenses without ransomware. Sec 66 D [19] deals with cheating by personation by utilizing computer resource, registered case is 2269 in 2017, 2704 in 2018, 5520 in 2019. In other sections of the IT Act, 1503 cases registered in 2017, 980 in 2018, 2720 in 2019. Sec 67 of IT [19] Act deals with publishing or transmitting obscene material in electronic form, registered case is 948



**Fig. 7** No. of cybercrimes registered under IT Act

in 2017, 1334 in 2018, 1845 in 2019. Sec 67A of IT Act [19] deals with publishing or transmitting of material containing sexually explicit act [19] in electronic form; the registered case is 401 in 2017, 1097 in 2018, and 1394 in 2019.

## ***4.2 Cybercrimes Registered Under IPC***

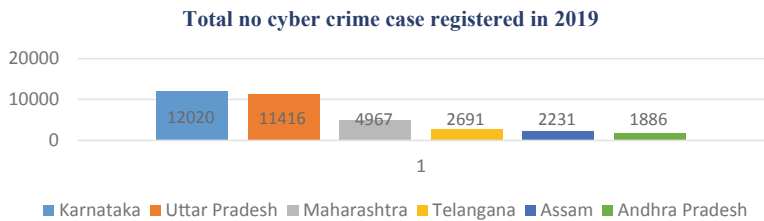
The cybercrime incidents under the Indian Penal code (IPC) [20] increased by 6.9% in 2018 as compared to 2017 (7976 cases registered in 2017 to 8500 cases in 2018), increased by 60.23% in 2019 as compared to 2019 (8500 cases in 2018 to 13,653 cases in 2019). Based on statistical analysis, top 8 crime heads registered under IPC as Cheating (Sec.420), ATMs, other offenses involving communication devices as medium or target, online banking fraud, other frauds under Sec.420 r/w Sec.465,468–471 IPC [20], cyberstalking/bullying of women/children (Sec.354D IPC), OTP frauds, cyberblackmailing/threatening (Sec.506,503,384 IPC) [20] 0.1993 cases registered in 2018 as linked to 3376 cases in 2019 under cheating (Sec 420) [20] of IPC. 2066 cases recorded in 2019 as related to 1543 cases in 2018 under ATM fraud under IPC. 1829 cases recorded in 2019 as related to 1680 cases in 2018 under other offenses involve communication devices as medium or target under IPC. A total no. of 2091 cases registered in 2019, compared to 965 cases in 2018 under online banking fraud under IPC. Other frauds under (Sec.420 r/w Sec.465,468–471) IPC [20] and total no. of cases are recorded 1151 in 2019 as related to 471 cases in 2018. Under cyberstalking/bullying of women/children (Sec.354D IPC [20]), the registered 752 cases in 2019 are compared to 706 cases in 2018 and 549 cases registered in 2019 are compared to 334 cases in 2018 under OTP fraud under IPC. Under cyberblackmailing/threatening (Sec.506,503,384 IPC) [20], a total no. of 266 cases registered in 2019 is compared to 311 cases registered in 2018.

## ***4.3 Statewise Crime Pattern Analysis***

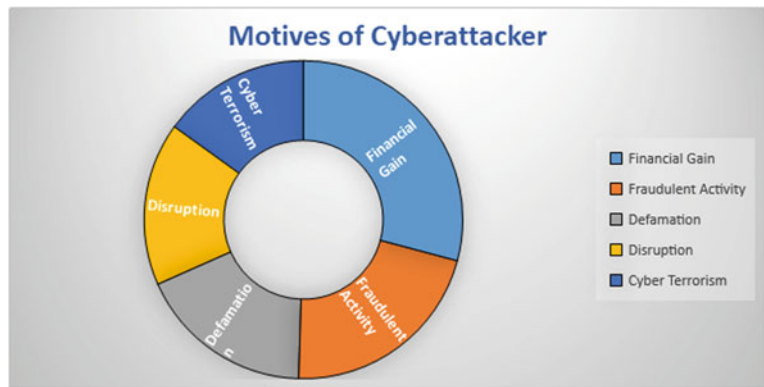
Karnataka ranked first with 12,020 registered cases in 2019 on the total number of cybercrime, followed by Uttar Pradesh with 11,416 registered cases, and Maharashtra with 4967 registered cases, represented in Fig. 8.

## **5 Motives of Attacker and Users not Reporting Cyberincidents**

As per the survey on “Cyber Crime Survey in 2017” organized by KPMG, there are five main motives of cyberattackers such as monetary gain, fraudulent activity,



**Fig. 8** Total no. cybercrime registered in 2019, top six states



**Fig. 9** Key motives of cyberattacker

defamation, disruption, cyberterrorism [21]. Figure 9 shows the key motives of the cyberattackers.

In reference to [21], owing to the inadequacy of understanding, 46% of end-users inside an organization do not report cybercrime incidents while 34% do not report such incidences for fear of penalties. 11 percent and 9% of cybercrimes go unreported for negative publicity and negligence in cyberreporting, respectively.

## 6 Initiative and User Awareness Program

The Government of India started various schemes to build understanding about cyber-crimes, e.g., sending short service message service to users periodically, organizing workshops, online security tips to educate users [2], setting up cybervolunteers [22], mock drills, helpline numbers, online reporting, forensic center, etc. The National Cyber Security Strategy (2020), designed to protect business-sensitive data, critical information infrastructure, affects national security and the economy. Different initiatives are undertaken taken up by the government, like setting up the Indian computer



emergency response team (CERT-In, [23]) and Cyber Suraksha Bharat [24]. The government picked up the initiative to set up National Critical Information Infrastructure Protection Center [25] and conducted Web site audit. They have also formed crisis management plans, malware protection [National Cyber Coordination Center] [26], etc.

## 7 Conclusion

Based on a statistical analysis of cybercrime data over the years, we estimated the upcoming trends of cybercrimes in the future. Such a forecast aids the government and various organizations in preparing better against cybercrimes. For a highly populated country like India, such preemptive measures can prevent a significant amount of data and monetary loss. User awareness, response to various initiatives started by the government, immediate reporting on time, law enforcement from time to time, advancing cyberforensic facilities, speed up the investigation process are the most significant factors toward mitigation of forthcoming cybercrimes. In the future, we plan to devise various mitigation strategies that would aid the government and concerned security personnel to prepare for the forecasted proportion of cybercrimes beforehand.

**Acknowledgements** Our sincere thanks to the National Crime Records Bureau (Ministry of Home Affairs), Government of India, for publishing datasets of cybercrime records.

## References

1. Statista. <https://www.statista.com/statistics/255146/number-of-internet-users-in-india/>
2. National Cyber Crime Reporting Portal, Government of India. <https://cybercrime.gov.in/>
3. The Economic Times. <https://cio.economictimes.indiatimes.com/news/digital-security/india-ranks-3rd-among-nations-facing-most-cyber-threats-symantec/63621655>
4. The Hindu. <https://www.thehindu.com/data/cyber-crime-cases-in-india-jumped-77-in-2017-compared-to-2016/article29889061.ece>
5. Information Technology Act 2000, Government of India (2000). <https://www.meity.gov.in/content/rules-information-technology-act-2000>
6. Shumway RH, Stoffer DS (2000) Time series regression and ARIMA models. In: Time series analysis and its applications, Springer, New York, NY, pp 89–212. ISBN: 978–1–4757–3261–0. <https://doi.org/10.1007/978-3-319-52452-8>.
7. National Crime Records Bureau (Ministry of Home Affairs), Government of India. <https://ncrb.gov.in/>
8. Sreehari A, Abinanth KJ, Sujith B, Unnikuttan PS (2018) A study of awareness of cyber crime among college students with special reference to Kochi. *Int J Pure Appl Math* 119(16):1353–1360. ISSN: 1314–3395
9. Chatterjee S, Kar AK, Dwivedi YK, Kizgin H (2019) Prevention of cybercrimes in smart cities of India: from a citizen's perspective. *Info Technol People* <https://doi.org/10.1108/ITP-05-2018-0251>, ISSN: 0959–3845

10. Gunjan VK, Kumar A, Avdhanam S (2013, September) A survey of cyber crime in India. In: 2013 15th international conference on advanced computing technologies (ICACT), IEEE, pp 1–6. <https://doi.org/10.1109/ICACT.2013.6710503>
11. Singh SK, Rastogi N (2018, February) Role of cyber cell to handle cyber crime within the public and private sector: an indian case study. In: 2018 3rd international conference on internet of things: smart innovation and usages (IoT-SIU), IEEE, pp 1–6. <https://doi.org/10.1109/IoT-SIU.2018.8519884>
12. Kshetri N (2016) Cybercrime and cybersecurity in India: causes, consequences and implications for the future. *Crime Law Soc Chang* 66(3):313–338. <https://doi.org/10.1007/s10611-016-9629-3>
13. Kumar PV (2016, March) Growing cyber crimes in India: a survey. In: 2016 International conference on data mining and advanced computing (SAPIENCE). IEEE, pp 246–251. <https://doi.org/10.1109/SAPIENCE.2016.7684146>
14. Datta P, Panda SN, Tanwar S, Kaushal RK (2020, March) A technical review report on cyber crimes in India. In: 2020 international conference on emerging smart computing and informatics (ESCI), IEEE, pp 269–275. <https://doi.org/10.1109/ESCI48226.2020.9167567>
15. Ministry of Electronics and Information Technology. <https://www.meity.gov.in/content/notification-it-act-2000>
16. Dolado JJ, Gonzalo J, Mayoral L (2002) A fractional Dickey-Fuller test for unit roots. *Econometrica* 70(5):1963–2006. <https://doi.org/10.1111/1468-0262.00359>
17. Duke University. <https://people.duke.edu/~rnau/411arim.htm>
18. The National Institute of Standard and Technology, Information Technology Laboratory. <https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4481.htm>
19. The information Technology Act, (2000). <https://www.indiacode.nic.in/bitstream/123456789/1999/3/A2000-21.pdf>
20. The Indian Penal Code. <https://legislative.gov.in/sites/default/files/A1860-45.pdf>
21. Cybercrime Survey Report 2017, KPMG India (2017). <https://home.kpmg/in/en/home/insights/2017/12/cybercrime-cybersecurity-law-enforcement-agencies.html>
22. National Cyber Crime Reporting Porta. <https://www.cybercrime.gov.in/>
23. Indian Computer Emergency Response Team. <https://www.cert-in.org.in/>
24. Cyber Surakshit Bharat Programme. <https://www.meity.gov.in/cyber-surakshit-bharat-programme>
25. National Critical Information Infrascture Protection center. <https://nciiipc.gov.in/>
26. National Cyber Crime Coordination Center. [https://www.mha.gov.in/division\\_of\\_mha/cyber-and-information-security-cis-division/Details-about-Indian-Cybercrime-Coordination-Centre-I4C-Scheme](https://www.mha.gov.in/division_of_mha/cyber-and-information-security-cis-division/Details-about-Indian-Cybercrime-Coordination-Centre-I4C-Scheme)

# Intelligent Indian Sign Language Recognition Systems: A Critical Review



Soumen Das, Saroj Kr. Biswas, Manomita Chakraborty,  
and Biswajit Purkayastha

**Abstract** Sign language (SL) is an alternative way of communication for hearing/speech impaired people. SL varies from country to country. Thus, different SLs came into existence such as American Sign Language (ASL), Chinese Sign Language (CSL), British Sign Language (BSL) and so on. In India, people use Indian Sign Language (ISL) as an alternative communication medium. Approximately, 2.7 million Indian people are deaf/dumb, out of which 98% population uses ISL for communication. But human interpreters are not available or affordable for identifying ISL. Thus, researchers are focussing on Sign Language Recognition System (SLRS) for the identification of ISL. But designing an SLRS for ISL is very difficult in comparison with other SL. ISL is very complex as it consists of single- and double-handed gesture and has an extensive vocabulary with similar gestures. This paper aims to provide a thorough study of ISL, its syntax/vocabulary and different trending techniques for designing an ISL recognition system. This paper also discusses various feature extraction techniques, classification methods used in ISL recognition and points out some challenges researchers face while developing an ISL recognition system.

**Keywords** Intelligent system · Machine learning · CNN · Sign language · ISL · Expert system · Challenges · SLRS · ISL syntax · Interpreters · Trending techniques

## 1 Introduction

Sign language (SL) is an alternative way of communication for hearing/speech impaired people. SL is a nonverbal language that visually transmits pattern using

---

S. Das (✉) · S. Kr. Biswas · M. Chakraborty · B. Purkayastha  
NIT Silchar, Fakiratilla, Silchar, Assam 788010, India  
e-mail: [Soumen\\_rs@cse.nits.ac.in](mailto:Soumen_rs@cse.nits.ac.in)

S. Kr. Biswas  
e-mail: [saroj@cse.nits.ac.in](mailto:saroj@cse.nits.ac.in)

many body parts, including manual elements like hand position and finger position as well as it has non-manual aspects like facial expression and lip movement. The important features for distinguishing SL are articulation points such as fingers joint, wrist joint, elbow joint and shoulder joint. It also includes different movement types such as forward, backward, left, right and facial expressions like glad, furious, confused and lips movement. SL varies from country to country. Thus, different SLs came into existence such as American Sign Language (ASL), Chinese Sign Language (CSL), British Sign Language (BSL), Indian Sign Language (ISL) and so on. ISL is unique and more complex than other SLs because it consists of single as well as double-handed gestures, and it has a large vocabulary of similar gestures which makes researchers' job more challenging to design a sign language recognition system (SLRS) for ISL recognition. In traditional approaches, a human interpreter is required to act as a mediator or a bridge between hearing/speech impaired people and normal people to reduce the communication gap. A competent interpreter must be aware of the language used by customers and their communication needs. The job of the translator has evolved throughout time [1] from "Helper" to "volunteer" to "communication facilitator" to "cultural and linguistic mediator". An interpreting work involves a variety of participants, each with its own set of requirements. So, a well-trained interpreter is required for this job. Human interpreters, on the other hand, are not always available and are also not cost effective. A human interpreter as a mediator undermines confidentiality and accuracy. Thus, human interpreters are not trustworthy where SLRS can serve as an effective and efficient substitute for a human interpreter in bridging the communication gap between hearing/speech impaired people and normal people.

In India, approximately 2.7 million people are deaf/dumb [2], out of which 98% population uses ISL for communication, so there is a huge communication gap between normal people and deaf/ dumb people and human interpreters cannot resolve the issue due to its unavailability and unaffordability. Thus, researchers are focussing on SLRS for the identification of ISL. Official ISL dictionary contains around 6,000 words, and new words are getting updated day by day. With such a large vocabulary, designing an SLRS is a challenging task. Since the last decade, various methods have been proposed for ISL recognition. Broadly, these are categorised into a sensor-based approach and a vision-based approach. In a sensor-based approach, data is collected by using different hardware devices and motion sensors such as data glove [3], inertial measurement unit (IMU) and leap motion controller (LMC). The sensor-based approach restricts signers' hand movement while performing any sign, and signers need to carry sensor devices that make this approach unsuitable for real-time application. Whereas in the vision-based approach, SLRS uses optical devices for the data collection part where signers do not require to carry or use any hardware devices. Moreover, signers allowed to perform signs freely. Thus, most recent works focus on a vision-based approach. For the recognition problem, SLRS employs a variety of machine learning and deep learning approaches. In ML algorithm, features extracted through digital image processing (DIP) techniques and morphological operation are given as input in various classifiers, such as vector machine (SVM) [4, 5], Bayes classifier and K-nearest neighbour (KNN) [6, 7], for the classification task. Based

on different signs, SLs are categorised into static sign and continuous sign. Static signs consist of a single gesture, whereas continuous signs consist of sequences of gesture. The ML algorithm performs well for static sign recognition. However, it fails to perform well for continuous sign recognition as continuous sign language recognition involves sequential data. Since last decade, DL technology has improved rapidly and most of the computer-vision-related problems are solved using this technique. Convolutional neural network (CNN), long short-term memory (LSTM) and recurrent neural network (RNN) are frequently used DL algorithms for sign language recognition [8, 9]. CNN is a special kind of DL algorithm where images are fed to the network directly. CNN can be used as a classifier and a feature extractor [10, 11], which makes this algorithm more suitable than ML algorithms as there is no need for manual feature extraction.

The paper's major goals are to (i) describe SL and introduce Indian Sign Language (ISL), (ii) provide an outline of ISL syntax and vocabulary, (iii) to talk about current ISL recognition and sentence creation strategies and (iv) to talk about ISL recognition problems.

## 2 Syntax/vocabulary of ISL

The first linguistic research on ISL initiated in 1978, and researchers observed that it is a natural language (NL) with its own syntax, grammar and morphology. ISL follows its own grammar, which is similar to spoken language. For sentence formation, ISL follows a basic time  $\rightarrow$  SOV (subject, object, verb) structure. In different parts of the country, ISL follows the same grammar rule [12]. In ISL, the first person (I, me, we, us) is always identified as a speaker, and the second person (you, your, yours) is identified as the addressing person. ISL does not use verbs such as is, am and are and article such as a, an, the and it only follows the root words. In ISL, numbers are translated by converting each digit of the number into a corresponding sign, sequencing from left to right. For example, 2021 will be represented as two-zero-two-one.

ISL vocabulary consists of a variety of words and a different way to express those words. There are some specific signs to represent gender in ISL. A finger marked above the lip represents the "male" sign, while the "female" sign is represented by a finger pointed to the nose. ISL consists of isolated as well as continuous signs, which includes both single and double-handed gesture. Signs such as a table, garden, mat and bench are represented by a double-handed gesture. In ISL, two or more signs are combined to represent a relationship. The male sign followed by the child signs express the "Son", whereas the "wife" is expressed by the female sign followed by the married sign [13]. ISL uses non-manual features such as facial expression to represent question mark and negation. To express question, eyebrows are raised with hand, whereas negation is represented by simple head movement. Negation is always used at the end of the sentence to represent a negative sentence, like "do not call me" represented as "call me not". ISL is at a developing stage, so every word

used in spoken language does not have an equivalent sign in the ISL dictionary, such as people name, places and books. For expressing these spoken words, fingerspelling is used. In ISL, single gestures can have different meanings based on their usage in different contexts.

### **3 Trending Techniques for ISL Recognition and Sentence Formation**

Researchers have worked on various SLRS for ISL recognition using machine learning (ML) and deep learning (DL) techniques. Based on the previous works, SLRS is broadly categorised into sensor-based approach and vision-based approach. Some of the important and trending techniques for ISL recognition are discussed in this section.

#### ***3.1 Sensor-Based Approach***

In the sensor-based approach, various sensors such as leap motion controller (LMC), Polhemus tracker, inertial measurement unit (IMU), gyroscope and accelerometer are used to detect hand movement, velocity and hand orientation. In this approach, feature extraction is not an essential step as data collected with the help of different sensors are sufficient for gesture recognition.

Gupta et al. [14] have proposed an SLRS for continuous ISL recognition, and they have acquired a sensor-based approach for the data collection part. In this process, an IMU is placed on the signer's hand which gives the tracking information of the hand while performing any sign. A DL technique is employed for the gesture recognition task. A modified t-LeNet architecture is proposed for the classification of a sign gesture. The proposed method is compared with two frequently used models: multi-channel deep convolution neural network (MC-DCNN) and t-LeNet for analysing time series data or sequential data. The proposed model has achieved an overall accuracy of 81.62%, which is comparatively better than t-Lenet (79.70%). MC-DCNN has achieved comparatively higher accuracy (83.94%), but the proposed model has achieved similar accuracy with fewer training parameters, and the training time is much lesser for the proposed method compared to MC-DCNN.

Heera et al. [15] have proposed a sensor-based approach for ISL recognition. The proposed method uses a data glove that keeps track of three kinds of information with the help of different sensors. A flex sensor is used to get finger bend information, a gyroscope is used to calculate the angular movement of hand whereas an accelerometer is used to get the acceleration of hand. A separate classification or feature extraction algorithm is not required for the proposed method as data collected through different

sensors are directly mapped with the individual signs. The collected data is transmitted to an android application with an SQLite database where all the information is stored through a Bluetooth module of the Arduino board. The proposed system converts the input sign into speech.

Deora et al. [16] have proposed an SLRS for recognition of static ISL based on colour gloves. The proposed method uses red and blue colour gloves for simplifying the segmentation process. The proposed method uses a data base that contains around 510 images of 25 alphabets and 9 numbers. The proposed system follows a series of steps for sign recognition. The proposed steps are segmentation, fingertip detection and PCA. The proposed system has achieved an accuracy of 94%.

Divya et al. [17] have introduced an electromyography (EMG)-based SLRS for ISL recognition. Muscles generates electric potential which is captured using EMG. ISL data is collected by using BIOPAC MP-45 from the signers. Data collected using this process contains external noise which is removed using a bandpass filter. Different features like simple square integral (SSI), mean absolute value (MAV), root mean square (RMS), standard deviation (STD), average amplitude change (AAC), maximum (MAX) and minimum (MIN) are extracted from an EMG data. For classification, SVM is used. The proposed system is tested on commonly used gesture like “Hello”, “Please”, “Good morning”, “Good night” and “Sorry”. The system has obtained an accuracy of 90%.

Chavan [18] have proposed an analytical ISL recognition model using LMC. The proposed model has a unique feature of classifying non-useful frames and eliminating them. Features related to hand, finger and palm are stored in a database using LMC. The test images features are also extracted using LMC and compared for recognition. The proposed architecture has two classification module, namely blocklist classifier and feature classifier. The blocklist classifier uses the random forest (RF) classifier to eliminate the non-useful frames. The feature classifier helps to recognize the sign gesture based on extracted features. After successful recognition of frames, it is fed to the word correction engine to get meaningful words.

Although the sensor-based approach achieves good accuracy, it comes with few limitations, such as signers need to wear sensor devices that do not allow signers to perform a sign freely. The sensor-based approach is not cost-effective, and country like India, where people cannot afford it, making this approach unsuitable for real-time use.

### **3.2 Vision-Based Approach**

In the vision-based approach, signers do not require any hardware or sensor device that causes difficulty while performing any sign. Thus, researchers are more focussing on this approach. In this approach, data is acquired through an optical device that can be a digital camera, webcam or cell phone camera. Over a period of time, researchers have proposed many vision-based techniques for recognizing ISL. Some of the important techniques are discussed here.

Tripathi et al. [19] have proposed a framework for continuous ISL recognition based on an ML algorithm. The proposed model takes videos of signs as input and gives the corresponding labels as output. Identifying signs from a continuous gesture is a difficult task. A gradient-based keyframe extraction method is used to split the continuous gesture into a sequence of isolated frames to solve this particular problem. Each isolated RGB frame is converted into HSV (hue, saturation, value), and background subtraction is done in the pre-processing step, followed by a feature extraction process. The orientation histogram is applied to extract essential features from the gesture sequence, and principle component analysis (PCA) is applied to reduce the dimension of extracted features. Different distance-based matrices such as chessboard distance, Euclidian distance, city block distance, correlation distance and cosine distance are used for the classification part. The experiment is performed on ten different types of sentence, where each sentence consists of two or more than two gestures. The proposed method gives a satisfactory result with Euclidian distance and correlation distance.

Sridhar et al. [20] also have proposed a DL model for ISL recognition. In the proposed experiment, authors have created their own database with frequently used ISL words as there is no standard data set available for ISL. The proposed data set contains around 4,287 videos of 263 sign word in 15 different categories of words. They have created the proposed dataset with the help of experienced signers. In the proposed experiment, several feature extraction techniques are combined with DL models and the proposed method tested with a subset of the original data set which has achieved an overall accuracy of 94.5%.

Rao et al. [21] have introduced a real-time application for continuous ISL recognition. The proposed system takes input from the signer by a selfie camera which gives the signer advantage of correcting themselves while performing a sign. The system goes through several pre-processing steps before the feature extraction process. A 2D gradient-based operator is used that reduces the background variation at different lighting conditions which helps to extract the head and hand portion from the sign images. The proposed method uses discrete cosine transformation (DTC) and PCA for feature extraction and for classification, different distance-based classifiers such as improved Euclidian distance, Euclidian distance and Mahalanobis distance classifier are used. The proposed system is tested on meaningful sentences. The performance of the proposed system is measured by using Word Matching Score (WMS). WMS is the ratio between the correctly identified sample to the total number of testing sample. Where Mahalanobis distance classifier has performed better than the other classifiers, with an average WMS of 85.88%.

Itkarkar et al. [22] have proposed an SLRS for ISL which takes input from the webcam and gives output in real-time. The proposed system can recognize alphabets and digits. It first takes input from the webcam and capture frames in YUY2 format where first Y represent luminance, U represents colour difference and the second Y also represents luminance. After that, it converts the frames into RGB format and then grayscale. The hand region is extracted from the images and after that, features are extracted using counter and convex hull method. The proposed system uses Harris



correlate algorithm for classification purpose. The proposed system is trained with 60 images for each sign and achieved an overall accuracy of 81.66%.

Sahoo et al. [23] have proposed an ISL recognition system for recognizing alphabets, numbers and words that consist of static and dynamic gesture. The proposed work has used a data set containing alphabets, numbers and a limited number of words which is used to define computer components in ISL. The data set contains single and double-handed gesture which represents ISL alphabets. The pre-processing step includes image restructuring and skin colour detection. Local histogram, direct pixel value (DPV) and hierarchical centroid (HC) techniques are used for feature extraction. Finally, K-nearest neighbour, naïve Bayes and neural network classifier are deployed to obtain the classification results. The proposed algorithm is applied to different categories of data separately to obtain the performance of the SLRS. The proposed experiment shows that the neural network classifier gives better accuracy than other classifiers with the same feature vectors.

Badhe et al. [24] have proposed a vision-based approach for ISL recognition. The vision-based system uses around 500 videos of 10 different sign performed by 10 signers. For feature extraction, the proposed method uses series of operations which includes frame extraction, background subtraction, thresholding and skin colour segmentation; after that, 28 furrier descriptors and vector quantization are used for feature extraction. For classification, a neural network-based model is proposed which includes three hidden layers with ReLu activation function. The proposed model trained with 300 videos for 50 epochs. The proposed system has obtained a training accuracy of 98% and validation accuracy of 63%.

Jayadeep et al. [25] have proposed an SRLS for recognizing ISL words commonly used in bank and daily life. The data set contains 1100 videos of banking words such as loan, deposit, credit, balance and debit card and everyday used words like come, sleep, food, home, love etc. Frames are extracted from the input videos in the first step, after which the data set is divided into training and test sets in 80:20 ratios. For feature extraction, InceptionV3 model is used. The extracted features from the previous layer are taken as input to the LSTM for classifying the input signs. The LSTM does not require manual feature engineering which makes this model suitable for the proposed task. The proposed model has reached an accuracy of 85% for banking words, whereas 92% for the words used in daily life.

Bhagat et al. [26] have proposed two separate models for recognizing static and dynamic word related to ISL. For the proposed experiment, a Microsoft RGB-D camera is used which helps to capture depth information of the sign images. For static gesture-based word recognition, a CNN model is proposed. The proposed model has trained with RGB and RGB depth images simultaneously. The training data set contains 45,000 depth images and 45,000 normal RGB images of sign alphabets related to ISL. The proposed model has achieved an accuracy of 99.6%. For dynamic gesture-based word recognition, an LSTM model is proposed. The model has trained with ten frequently used words like Wi-Fi, lock, maths, mall, low, win, machine, local, license and aeroplane. The proposed dynamic gesture recognition model has achieved an accuracy of 99.08%.

SL recognition is not limited to just recognizing alphabets, numbers and words. In real time, the SLRS needs to generate sentence from the recognized sign, and for this purpose, SLRS needs to identify sign words and find the context between words to generate a meaningful sentence. In recent years, researchers has done work in this direction [27–29].

In the last decade, researchers have proposed different feature extraction techniques and classification techniques for ISL recognition. Following Table 1 has summarised different SLRS based on SL type, feature extraction method, classification method and performance.

## 4 Challenges for ISL Recognition

SL recognition itself is a challenging task, whereas ISL proposes some new challenges which makes researchers job more difficult for designing an SLRS. For language recognition, identification of grammar rules of that language is necessary, but in the case of ISL, the grammar rules are not adequately documented due to which making an SLRS for ISL recognition is a challenging task for the researchers. As discussed in the previous section, the linguistic research on ISL started in 1978, so standardization of ISL is still under process. Thus, ISL varies in different parts of India, which makes the researcher's job difficult to make a uniform SLRS for ISL recognition. The unavailability of a standard dataset with a large vocabulary is another problem for ISL recognition as developing an SLRS needs adequate and authentic data. Unlike other SL, ISL alphabets and words use single-handed as well as double-handed gestures. While posing double-handed gestures, sometimes signers' hands overlap, making the segmentation and recognition task more complex than other sign languages.

## 5 Conclusion

This paper gives a brief discussion about ISL and how ISL is different from other existing SL. This paper discusses the syntax or grammar rule and vocabulary for ISL. The primary goal of the paper is to give an outline of the ISL and the trending techniques for ISL recognition. Based on previous works, this paper categorised the ISL recognition techniques into sensor-based approach and vision-based approach. Lastly, this paper summarises different approaches for feature extraction and classification techniques acquired by the researchers over the period. While designing an ISL recognition system, researcher faces different challenges; these are also discussed in this paper. Most of the research works in SLRS are based on either static sign recognition or isolated word recognition, and very few works have been done for sentence recognition or continuous sign recognition. Thus, continuous SL recognition and sentence formation can be a potential area of research in future.

**Table 1** Summary of feature extraction methods and classification methods used in SLRS

Author name	SL Type	Feature extraction method	Classification method	Accuracy/Performance
Athira et al. [30]	Static	Zernike moments	Multi class SVM	91%, 89%
Bhumika Gupta et al. [6]	Static	HOG, SIFT	KNN	78.84%, 80%, 90%
Kusumika Krori Dutta et al. [31]	Static	PCA	KNN	95.84%, 94.88%
Sajanraj et al. [11]	Static	CNN	R-CNN	99.56%, 97.26%
Ananya Choudhury et al. [32]	Continuous	Segmentation, hand tracking	Conditional random field (CRF)	90%, 86%
Raheja [33]	Static and continuous	Hu moments, hand tracking	Multi class SVM	97.5% for selected 4 signs
Mittal et al. [34]	Continuous	CNN	Modified LSTM	89.50% for sign word, 72.30% for signed sentence
Karush Suri et al. [35]	Continuous	IMU	CapsNet	94% for 3 routings, 92.5% for 5 routings
Likhar et al. [36]	Continuous	CNN	LSTM	99.08%
Jayadeep et al. [25]	Continuous	Inception V3	LSTM	85%, 92%

## References

1. Davis JE (2005) Working with sign language interpreters in human service settings in national organization for human service (NOHS). 25(1):41–52
2. Dudhal A, Mathkar H, Jain A, Kadam O, Shirole M (2018) Hybrid SIFT feature extraction approach for indian sign language recognition system based on CNN. In: Proceedings of the international conference on ISMAC in computational vision and bioengineering (ISMAC-CVB), Springer, pp 727–738
3. Shukor AZ, Miskon MF, Jamaluddin MH, Bin F (2015) A new data glove approach for malaysian sign language detection. Proc Comput Sci 76(1):60–67
4. Kumar A, Thankachan K, Dominic MM (2016) Sign language recognition. In: 3rd international conference on recent advances in information technology (RAIT), IEEE, Dhanbad, pp 422–428
5. Chandra MM, Rajkumar S, Kumar LS (2019) Sign languages to speech conversion prototype using the SVM classifier. In: IEEE Region 10 conference (TENCON), Kochi, pp 1803–1807
6. Gupta B, Shukla P, Mittal A (2016) K-nearest correlated neighbor classification for Indian sign language gesture recognition using feature fusion. In: International conference on computer communication and informatics (ICCCI), IEEE, Coimbatore

7. Sadik F, Subah MR, Dasider AG, Moon SA, Ahbab SS, Fattah SA (2019) Bangla sign language recognition with skin segmentation and binary masking. In: International WIE conference on electrical and computer engineering (WIECON-ECE), IEEE, Bangalore, pp 1–5
8. Aly S, Aly W (2020) DeepArSLR: a novel signer-independent deep learning framework for isolated arabic sign language gestures recognition. *IEEE Access* 8(1):83199–83212
9. Sreenivas A, Maheshwari M, Jain S, Choudhary S, Vadivu G (2020) Indian sign language communicator using convolutional neural network. *Int J Adv Sci Technol* 29(3):11015–11031
10. Hasan MM, Srizon AY, Mehedi Hasan MA (2020) Classification of Bengali sign language characters by applying a novel deep convolutional neural network. In: IEEE region 10 symposium (TENSYP), Dhaka, pp 1303–1306
11. Sajanraj TD, Beena MV (2018) Indian sign language numeral recognition using region of interest convolutional neural network. In: Second international conference on inventive communication and computational technologies (ICICCT), IEEE, Coimbatore, pp 636–640
12. Dasgupta T, Dandpat S, Basu A (2008) Prototype machine translation system from text to Indian sign language. In: Proceedings of the IJCNLP-08 workshop on NLP for less privileged languages, ACM, Hyderabad
13. Kumar SP, Kaur S (2020) Sign language generation system based on indian sign language grammar. *ACM Trans Asian and Low-Resour Lang Info Process* 19(4):1–26
14. Gupta R, Ranjan S (2020) Comparative analysis of convolution neural network models for continuous indian sign language classification. *Proc Comput Sci* 171(1):1542–1550
15. Yarisha Heera S, Murthy MK, Sravanti VS, Salvi S (2017) Talking Hands—an indian sign language to speech translating gloves. In: International conference on innovative mechanisms for industry applications (ICIMIA), IEEE, Bengaluru, pp 746–751
16. Deora D, Bajaj N (2012) Indian sign language recognition. In: 1st international conference on emerging technology trends in electronics, communication and networking, IEEE, Surat
17. Divya B, Delpha J, Badrinath S (2017) Public speaking words (Indian Sign Language) recognition using EMG. In: International conference on smart technologies for smart nation (SmartTechCon), IEEE, Bengaluru, pp 798–800
18. Chavan P, Ghorpade T, Padiya P (2016) Indian sign language to forecast text using leap motion sensor and RF classifier. In: Symposium on colossal data analysis and networking (CDAN), IEEE, Indore
19. Tripathi K, Baranwal N, Nandi GC (2015) Continuous Indian sign language gesture recognition and sentence formation. In: Eleventh international multi-conference on information processing (IMCIP), vol 54. Elsevier, Allahabad, pp 523–531
20. Sridhar A, Ganesan RG (2020) Include: a large scale dataset for indian sign language recognition. In: MM '20: Proceedings of the 28th ACM international conference on multimedia, ACM, USA, pp 1366–1375
21. Ananth Rao G, Kishore PVV (2018) Selfie video based continuous Indian sign language recognition system. *Ain Shams Eng J* 9(4):1929–1939
22. Itkarkar RR, Nandi A, Mane B (2017) Contour-based real-time hand gesture recognition for indian sign language. In: Computational intelligence in data mining, vol 556. Springer, Pune, pp 683–691
23. Sahoo AK, Sarangi PK, Goyal P (2020) Indian sign language recognition using soft computing techniques. *Machine Vision Inspection Syst* 1(1):37–65
24. Badhe PC, Kulkarni V (2020) Artificial neural network based indian sign language recognition using hand crafted features. In: 11th International conference on computing, communication and networking technologies (ICCCNT), IEEE, Kharagpur
25. Jayadeep G, Vishnupriya NV, Venugopal V, Vishnu S, Geetha M (2020) Mudra: convolutional neural network based indian sign language translator for banks. In: International conference on intelligent computing and control systems (ICICCS), IEEE, Madurai, pp 1228–1232
26. Bhagat NK, Vishnusai Y, Rathna GN (2019) Indian sign language gesture recognition using image processing and deep learning. In: Digital image computing: techniques and applications (DICTA), IEEE, Perth

27. Raghuveera T, Deepthi R, Mngalashri R, Akshaya R (2020) A depth-based Indian sign language recognition using microsoft kinect. *Indian Acad Sci* 46(2):33–46
28. Agarwal SR, Agrawal SB, Latif AM (2015) Sentence formation in NLP engine on the basis of indian sign language using hand gestures. *Int J Comput Appl* 116(17):18–22
29. Gangadia D, Chamaria V, Doshi V, Gandhi J (2020) Indian sign language interpretation and sentence formation. In: *IEEE international conference (PuneCon)*, Pune, pp 71–76
30. Athira PK, Sruthi CJ, Lijiya A (2019) A signer independent sign language recognition with co-articulation elimination from live videos: an indian scenario. *J King Saud Univer –Comput Infor Sci* 31(2):1–11
31. Dutta KK, Bellary SAS (2017) Machine learning techniques for indian sign language recognition. In: *International conference on current trends in computer, electrical, electronics and communication (ICCTCEEC)*, IEEE, Mysore, pp 333–336
32. Choudhury A, Talukdar AK, Sarma KK(2014) A conditional random field based indian sign language recognition system under complex background. In: *Fourth international conference on communication systems and network technologies*, IEEE, Bhopal, pp 900–904
33. Raheja JL, Mishra A, Chaudhary A (2016) Indian sign language recognition using SVM. *Pattern Recog Image Anal* 26(2):434–441
34. Mittal A, Kumar P, Roy PP, Balasubramanian R, Chaudhuri BB (2019) A modified-LSTM model for continuous sign language recognition using leap motion. *IEEE Sens J* 19(16):1–8
35. Suri K, Gupta R (2019) Continuous sign language recognition from wearable IMUs using deep capsule networks and game theory. *Comput Electri Eng* 78(1):493–503
36. Likhar P, Bhagat NK, Rathna GN (2020) Deep learning methods for indian sign language recognition. In: *10th International conference on consumer electronics (ICCE)*, IEEE, Berlin

# Performance Analysis of Gujarati Script Recognition Using Multiclass and Multilabel Classification



Shweta Dalal, Sarosh Dastoor, Keyur Zaveri, Virti Choksi, and Karvin Shah

**Abstract** In the era of digital world, the paper documents are turning into digitally accessible and editable documents. It has led to the advancement of the optical character recognition system. So to take the document work into digital format, the research on optical characters for Gujarati script has been carried out in this paper. We have tried to inculcate all the Gujarati consonants, vowels and their combinations for recognition and classification, which has not been done so far. The approach of deep learning and optimization had been used to train our network using multiclass and multilabel classification methods on a systematically generated data set. The chief goal of the research is to include all the Gujarati characters and analyse the performance of these two classification methods used for the character recognition of Gujarati language and produce comparative results in terms of validation accuracy, validation loss and number of epochs. The various models used for comparison over the two methods are DenseNet201, EfficientNetB0 and ResNet50.

**Keywords** OCR · Text recognition · Python · Deep learning · Classifiers · Gujarati script

## 1 Introduction

Recognizing characters of any language has been a challenging task in the field of deep learning and computer vision especially for Indian primitive scripts. Optical character recognition is a method to study patterns of the characters by training a neural network on them. Though there are several developments in text recognition nowadays, as such no concrete work has been done for the recognition of a native script like Gujarati which is the third most popular language in India.

In multiclass classification method, each sample is assigned a single label at a time. It is a one-vs-rest type of classifier which involves training a single classifier per class, with the samples of that class as positive samples and all other samples as

---

S. Dalal (✉) · S. Dastoor · K. Zaveri · V. Choksi · K. Shah  
Sarvajanik College of Engineering and Technology, Gujarat, India  
e-mail: [sarosh.dastoor@scet.ac.in](mailto:sarosh.dastoor@scet.ac.in)

negatives. This method becomes computationally expensive as the number of classes increases. So in such a case, multilabel classification can reduce this computational complexity and train our model much faster producing better results by predicting more than one non-exclusive class at a time. We have trained our models using the above mentioned methods, and the results which we obtained are then compared on the basis of validation accuracy and validation loss obtained. Our data set includes systematically generated Gujarati script images with 400 classes. In such a case, it becomes crucial to prove the superiority of one classifier over the another in terms of validation accuracy and computational complexity.

## 2 Background and Related Work

Automatic recognition and reading of characters have been tried and attempted in various languages. Also devices for automatic recognition of handwritten scripts are available in the market, but they are limited to only few languages. Yet, the recognition part of these scripts fails to include all its characters and styles because of various challenges like scene complexity, conditions of uneven lighting, fonts, multilingual environments and many more [1]. Gujarati script, on a similar note, includes more than 400 characters including vowels. Classifying such a huge no. of classes so that it can correctly recognize input characters is a challenging task

In [2], it is proposed that CNN gives better results and works more efficiently when we use it for extraction of features rather than classifying the data. On the same note, classification of 50 Bangla basic characters has been carried out in [3], and a comparative analysis for different architectures like Vgg-Net, DenseNet, ResNet and others has been made using the technique of multiclass classification. A similar data set with same number of classes has been used to train on Bangla script using SVM classifier which supports a multiclass classification process [2], and an accuracy of 98.37 % was obtained though it took much training effort without fine-tuning. But it fails to include all character combinations. A deep convolutional neural network with two convolutional layers, followed by three densely connected layers along with dropout regularization technique, is used to classify total 80 characters (11 vowels, 39 consonants, 24 compound letters and 10 numerals) of Bengali handwritten characters [4] in which resulted in 89.93 % accuracy. In [5], the Kannad characters are deformed and then trained upon using AlexNet over 497 classes. This fetched 91.3% accuracy for printed characters and 92% for handwritten characters. While, in [6] the highest accuracy of 91.63 % has been obtained using support vector machine (SVM) with radial basis function (RBF) kernel over 55 samples of Gujarati characters. But the drawback of this method is that it provides low accuracy for similar looking and confusing characters. In [7], it has been proved that DenseNet121 architecture gives the best accuracy out of various CNN architectures without proper fine-tuning over the systematically generated data set. But the common problem which seems to be faced is the training inefficiency which results as the size of data set increases which includes combination of characters (consonants+vowels).

As an alternative to this approach, using multilabel classification, the output vector length can be significantly reduced. It is used when a number of classes are large and possibly comparable with the number of observations, with one specific characteristic: a data point is tagged with several labels [8]. But in multilabel classification, assigning labels is difficult since each instance may belong to more than one class [9]. Multilabel classification method has been vividly used for predicting and classifying large number of classes. This approach has been explained and implemented in [10] which uses multilabel text classification where model is applied to predict the different product categories with data set manually labelled annotations about the categories associated with each text description. 0.86 precision and 0.81 recall have been obtained with 0.25 threshold. In [11], multilabel classification technique has been deployed to recognize Bengali characters by splitting the characters of a Bengali grapheme into three labels, i.e. grapheme roots, vowel diacritics and consonant diacritics. Thus, by using such a method and efficiently classifying all Gujarati script characters, an assistance system can be developed for visually impaired people for directing them towards reading and understanding Gujarati script [12].

Recognition of Gujarati characters still lags in inclusion of all vowels and its combination with characters. After concluding the limitation of multiclass classification, a concrete and an efficient algorithm with the help of multilabel classification which can recognize and classify all of those characters has been proposed in this paper.

### 3 Characteristics of Gujarati Script

This section shows the representation of Gujarat script. It is a native language used in the state of Gujarat in India, which is an Indo-Aryan-type language. The association of the vowels with the consonants makes it a challenging problem for recognition. Such a combination creates more than 400 such Gujarati characters. Some randomly chosen characters from the data set have the following phonetics, “sa”, “kha”, “ki”, “fu”, “ya”, “a”, “chi”, “jo”, “ha” and “shu”.

### 4 Methodology and Approach

We have trained various CNN models on Gujarati script using multilabel and multiclass classification techniques and have done a comparative analysis to prove the superiority of one over the other. The various models used for comparison over the two methods are EfficientNet, DenseNet201 and ResNet50.



4.1 Data Set

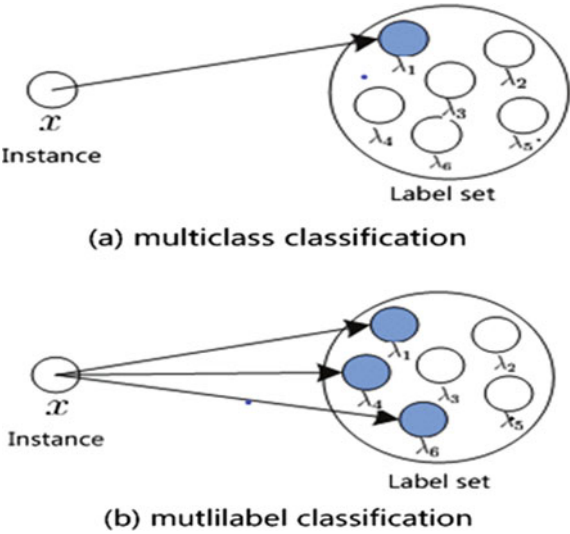
The data set is obtained from Indian Language Technology Proliferation and Deployment Centre. It is incorporated with 400 classes of Gujarati characters of which 80% goes to train set and remaining 20% is used for test set. The sample of data set containing some randomly chosen characters is shown in Fig. 1 given below.

Figure 2 illustrates the difference in the methodology of classification using multiclass and multilabel. In case of multiclass classification method, for any instance, say “ $x$ ”, only one label is assigned at a time. While in the case of multilabel classification, more than one labels can be assigned at a time for a single instance  $x$ . In other words, we predict the properties of a data point that are mutually exclusive. Thus, each data point can be assigned multiple classes at a time.



Fig. 1 Sample data set

Fig. 2 Comparison of multiclass and multilabel classification



4.2 Multiclass Classification

In multiclass classification task, each data point is given only one label. It is trained using a set of examples associated with just one class  $n$  of a set of classes  $N$ , where  $|N| > 1$ . If  $|N| = 2$ , then it is called binary classification. If  $|N| > 2$ , then it is said to be multiclass classification. Here each class is assigned each respective label. So the total number of labels sums up to total number of classes. In our case, the total output units are 400. The one-hot encodings corresponding to some of the classes are shown in Fig. 3 given below.

The models used for analysing the classification efficiency are EfficientNet, DenseNet201 and ResNet50. In the last layers of these architectures, sigmoid layer has been appended with 400 units since there are total 400 output classes. We have used TensorFlow and Keras libraries to implement these models. We have also included callbacks and fixed the number of training epochs for each architecture. Callbacks can be used to view on internal states and statistics of the model during training.

4.3 Multilabel Classification

This approach has not been so far explicitly used for character recognition. Here the data may belong to more than one class simultaneously. The advantage of splitting any character into consonant and vowel has been utilized in this algorithm. Example is shown in Fig. 4.

Thus, each character is a combination of consonant and vowel. Thus, each data character can be assigned one or more classes. There are total 400 classes. The advantage of this method is that the number of output units reduces to 56 units (34 consonants + 13 vowels + 9 conjunct characters) from 400. This can save a lot of computational power and training time since now the output one-hot encoding

Classes/Labels	ક	કલ	ખ	ખલ	ગ	ગલ	ઘ	ઘલ	ચ	ચલ
ક	1	0	0	0	0	0	0	0	0	0
કલ	0	1	0	0	0	0	0	0	0	0
ખ	0	0	1	0	0	0	0	0	0	0
ખલ	0	0	0	1	0	0	0	0	0	0
ગ	0	0	0	0	1	0	0	0	0	0
ગલ	0	0	0	0	0	1	0	0	0	0
ઘ	0	0	0	0	0	0	1	0	0	0
ઘલ	0	0	0	0	0	0	0	1	0	0
ચ	0	0	0	0	0	0	0	0	1	0
ચલ	0	0	0	0	0	0	0	0	0	1

Fig. 3 Multiclass classification labelling

કા = ક + આ	પૌ = પ + ઔ	લિ = લ + ઇ
ને = ન + એ	ગુ = ગ + ઉ	નં = ન + અં

Fig. 4 Character combination

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
labels	.ah	.am	.aha	an	.a	i	.ee	.u	.oo	.ru	.ay	e	.ai	.ov	.o	.ou	.r	.k	k
0 ક	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1 કે	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2 કઃ	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3 કઃ	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4 કઃ	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
5 કા	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
6 કિ	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
7 કી	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
8 કુ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
9 કૂ	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
10 કઃ	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
11 કે	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
12 કે	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
13 કે	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
14 કો	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
15 કો	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
16 કો	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
17 ક	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

Fig. 5 Multilabel classification labelling

consists of only 56 units since there are only 56 output classes. The output classes are named in English language for simplicity. The data can belong to one or more classes at the same time in multilabel classification. The encodings corresponding to some of the classes are shown in Fig. 5.

These labels are then imported into our framework, and the data set is trained onto it. Then we have then trained our model on EfficientNet, DenseNet201 and ResNet50 in TensorFlow. Sigmoid activation function is appended in the final layer of the architecture since the final score for each class should be independent of each other. We cannot apply softmax activation, because softmax converts the score into probabilities taking other scores into consideration. Sigmoid converts each score of the final node between 0 and 1 independent of what the other scores are. Since we are using a sigmoid activation function, we have to go with binary-cross-entropy loss.

**Table 1** Performance analysis of multiclass classification

Multiclass classification	Parameters	DenseNet201	EfficientBO	ResNet50
	Validation accuracy	0.81	0.52	0.3
	Validation loss	0.644	1.44	3.43
	Number of epochs	23	20	1.7

**Table 2** Performance analysis of multilabel classification

Multiclass classification	Parameters	DenseNet201	EfficientNetBO	ResNet50
	Validation F1_score	0.94	0.78	0.33
	Validation loss	0.44	1.97	3.24
	Number of epochs	30	27	20

## 5 Simulation and Results

Multiclass classification and multilabel classification have been implemented using Python, and TensorFlow framework was used for the training purpose. The performance of these two classification methods has been compared over Gujarati script, and the superiority of multilabel classification in terms accuracy has been proven owing to our results. For performance analysis of multiclass classification, parameters like validation accuracy and validation loss have been considered, whereas for multilabel classification we define a metric call “F-score” to monitor the performance of the model. In the case of multilabel classification, it is advised to use “F-score” instead of “accuracy” [13, 14]. F-score, which is mean of the precision and recall, is used to calculate the test accuracy. Precision is the number of true positive results divided by the number of all positive results, including those not identified correctly, and the recall is the number of true positives divided by the number of all samples that are supposed to be identified as positive. TensorFlow provides a built-in F-score metric in TensorFlow-Addons module. We have also used loss function related to F-score provided by TensorFlow-Addons modules (Tables 1 and 2).

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The given tables illustrate the comparison of two classification methods in terms of validation accuracy for multiclass classification and validation F1-score for multilabel classification. It can be seen that better results with more accuracy have been obtained for the three architectures using the method of multilabel classification.

DenseNet201 provided 13% more accuracy while EfficientNetB0 performed better by 26%. However, while training on ResNet50, not much change was observed in the accuracy score.

Figures 6 and 7 compare the validation loss against the number of epochs for different architectures which is considered an important parameter to analyse the performance. For DenseNet201, the validation loss plummets between first ten epochs and saturates at 0.44 in multilabel classification, while using multiclass classification the initial loss did not vary much and saturated at 0.644. In EfficientNetB0, when using multiclass classification, the loss showed a steep downfall before saturating to 1.44 loss value in 20 epochs, whereas the loss remained higher when trained using multilabel classification with the value settling at 1.97 without any further difference in 25 epochs. While using ResNet50, the validation loss remained above 3 with multilabel classification performing better. Moreover, the training time also reduced in multilabel classification since the output number of units is less.

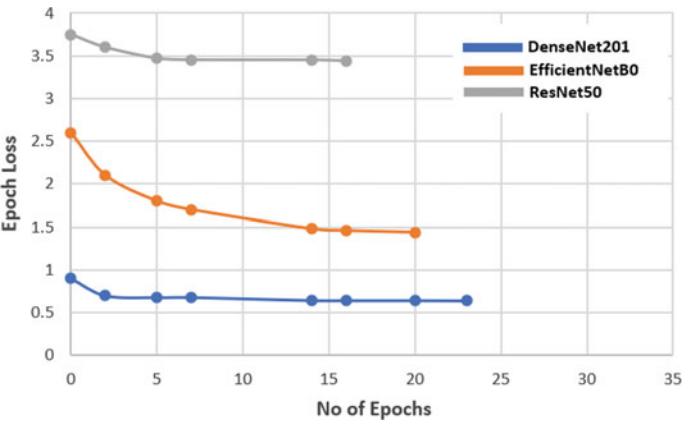


Fig. 6 Comparison graph of validation epoch loss for multiclass classification

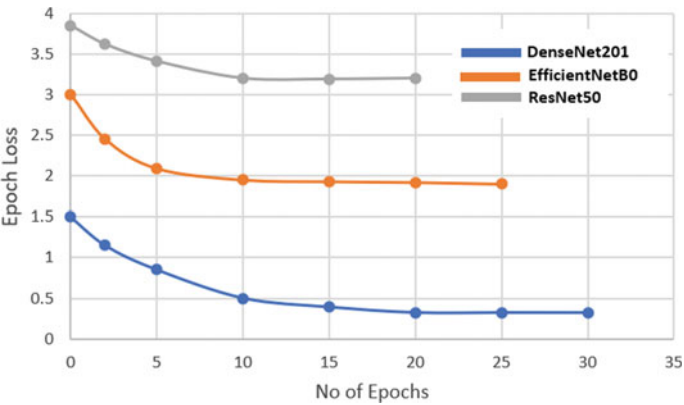


Fig. 7 Comparison graph of validation epoch loss for multilabel Classification

## 6 Conclusion

In this paper, two classification methods are compared for the recognition of Gujarati characters. The data set has been used to train on three CNN architectures, namely DenseNet201, EfficientNetB0 and ResNet50. These three architectures are used for multiclass and multilabel classification. Using ReduceLROnPlateau callback, the training was stopped when validation accuracy did not improve after three consecutive epochs. ReduceLROnPlateau callback reduces the learning rate when a metric has stopped improving. This callback monitors the accuracy, and if accuracy does not increase after a “patience” number of epochs, the learning rate is reduced. In our research, we have successfully trained model on all the Gujarati characters and concluded that due to the large data set of Gujarati script and huge number of classes, multiclass classification technique was not successful in converging the result to an optimized accuracy. As the no. of classes increases, the training process became slower and accuracy saturated at a very low value. On the contrary, all the three architectures performed better when classified using multilabel classification which uses the concept of assigning multiple classes to a single data simultaneously, thus reducing output number of units. This resulted in more training accuracy using comparatively less number of epochs though the validation loss flattened at a relatively high value in multilabel classification. Thus, we were able to recognize and classify all the characters of Gujarati efficiently. Also, DenseNet architecture showed the best performance in terms of accuracy as compared to other models in both the classification methods, and the reason it took more number of epochs was the high validation loss which did not flatten out early on.

## References

1. Hamad KA, Kaya M (2016) A detailed analysis of optical character recognition technology. *Int J Appl Math Electronics Comput* 4:244–249
2. Maitra DS et al (2015) CNN based common approach to handwritten character recognition of multiple scripts. In: 2015 13th International Conference on Document Analysis and Recognition (ICDAR), pp 1021–1025
3. Chowdhury R, Hossain M, Islam R, Andersson K, Hossain S (2019) Bangla handwritten character recognition using Convolutional Neural Network
4. Purkaystha B, Datta T, Islam MS (2017) Bengali handwritten character recognition using deep convolutional neural network. In: 2017 20th International Conference of Computer and Information Technology (ICCIT), pp 1–5. <https://doi.org/10.1109/ICCITECHN.2017.8281853>
5. Rani NS, Chandan N, Jain S, Kiran H (2018) Deformed character recognition using convolutional neural networks. *Int J Eng Technol* 7:1599. <https://doi.org/10.14419/ijet.v7i3.14053>
6. Naik VA, Desai AA (2017) Online handwritten Gujarati character recognition using SVM, MLP, and K-NN. In: 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp 1–6. <https://doi.org/10.1109/ICCCNT.2017.8203926>
7. Performance Analysis of Various Trained CNN Models on Gujarati Script [https://doi.org/10.1007/978-981-15-7062-9\\_48](https://doi.org/10.1007/978-981-15-7062-9_48)
8. Tsoumakas G, Katakis I (2007) Multi-label classification: an overview. *Int J Data Warehousing Mining (IJDWM)*, IGI Global 3(3):1–13

9. Tahir MA, Kittler J, Bouridane A (2012) Multilabel classification using heterogeneous ensemble of multi-label classifiers. *Pattern Recogn Lett* 33:513–523
10. Arroyo R, Jimenez D, Martínez-Cebrián J (2020) Multi-label classification of promotions in digital leaflets using textual and visual information
11. Alam S, Reasat T, Sushmit AS, Siddiquee SM, Rahman F, Hasan M, Humayun AI (2020) A large multi-target dataset of common bengali handwritten graphemes
12. Shah, Bandariya J, Jain G, Ghevariya M, Dastoor S (2019) Cnnbased auto-assistance system as a boon for directing visually impaired person. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp 235–240
13. Pillai I, Fumera G, Roli F (2012) F-measure optimisation in multi-label classifiers, pp 2424–2427
14. Lipton ZC, Elkan C, Narayanaswamy B (2013) F1-optimal thresh-olding in the multi-label setting. In: The Computing Research Repository (CoRR)

# Real-Time Multi-Object Tracking of Pedestrians in a Video Using Convolution Neural Network and Deep SORT



S. M. Praveenkumar, Prakashgouda Patil, and P. S. Hiremath

**Abstract** Multi-Object tracking in video processing is a challenging task in computer vision. Typically, tracking the objects in a video is performed using binocular vision or top-down camera movements. These methods encounter the problem of rigid boundaries on the counting area application scene. To overcome this problem, in this paper, a novel object detection algorithm is proposed by revamping YOLOv3 and real-time multi-object tracking is done by employing Deep SORT to track the targets by using movement representation and data association algorithms. It is a tracking-by-detection algorithm. The proposed algorithm is experimented on different public datasets and own dataset of videos of real scenes relevant to pedestrian movements. The results of experiments indicate that the proposed algorithm yields higher accuracy and exhibits better performance in real-time object tracking as compared to other object tracking techniques in the literature and hence is a reliable object-tracking algorithm.

**Keywords** Convolutional neural network · Video processing · Multi object tracking · Data association · Deep SORT

## 1 Introduction

Multiple Object Tracking (MOT) is the process of determining the number and states of various dynamic objects successively based on noisy sensor measurements. Object tracking in a video is a crucial technology with potential applications in many areas, e.g. robotics, video surveillance, autonomous driving, traffic control, industry automation, telemedicine, and sensor networks [1, 2]. In recent years, many

---

S. M. Praveenkumar (✉) · P. Patil · P. S. Hiremath  
KLE Technological University, Hubballi, Karnataka 580031, India  
e-mail: [Praveenkumar\\_sm@kletech.ac.in](mailto:Praveenkumar_sm@kletech.ac.in)

P. Patil  
e-mail: [prakashpatil@kletech.ac.in](mailto:prakashpatil@kletech.ac.in)

P. S. Hiremath  
e-mail: [pshiremath@kletech.ac.in](mailto:pshiremath@kletech.ac.in)



researchers in academics as well as industries are engaged in research work related to object tracking algorithms [2–4].

The initial development of automated approaches for monitoring pedestrian traffic is based on pressure sensors or photoelectric sensor. But, such methods are found to be not accurate enough in case of flow of higher pedestrian density with accompanying occlusions. Of late, pedestrian traffic monitoring methods making use of vertical downward stereo cameras have been developed [4, 5], which are good enough for obtaining pedestrian traffic statistics at present. This is on account of the fact that the pedestrian occlusions can be avoided by using vertical downward camera and the complex color background can be filtered by the binocular-vision-based 3D reconstruction algorithm. However, its usage suffers from severe limitation in terms of the site and visual field of the camera. Hence, it is appropriate in cases of indoor scenes with narrow entrances, but not to outdoor scenes with wide entrances often seen in commercial streets. However, it is not economically viable to deploy front-down monocular cameras in large numbers covering the entire city.

The tracking-by-detection framework is a popular approach to overcome this problem, wherein a detector localizes objects in each image of the frame sequence and then associates the detected objects in the succeeding image frames. However, the limitations of the object detector lead to false positives and missing detections. Also, due to the challenges posed by frequent occlusions and large variations in illumination, target scale, fog concentrations, etc., any algorithm based on front-down monocular camera for determining pedestrian flow statistics requires more accurate detector and tracker. In real time object tracking, the objects are found to be occluded at some instances and not occluded at other instances as time passes. Similarly, the difficulty can arise due to illumination and scale variations of the same object in the process of the object being tracked. The design of efficient detection and tracking algorithms, which accounts for such difficulties, becomes a major research goal. Due to the recent development of deep learning algorithms [6–11], pedestrian detection is achieved more accurately than conventional approaches. Certainly, the detection results of higher accuracy lead to better performance [12], which helps to design algorithm for obtaining pedestrian flow statistics by using the front-down monocular camera with greater precision.

In this paper, a novel algorithm for multi-object detection and tracking in a video is proposed. It is a tracking-by-detection algorithm. Firstly, a revamped YOLOv3 is used to detect object appearing in the camera view and, secondly, Deep SORT is applied to track the targets by using movement representation and data association algorithms. YOLOv3 is convolution neural network (CNN), in which input image is transformed to feature map through the convolution operation and then classified. In traditional ANN, the neural network accepts features as inputs and then does the classification. Deep SORT is an extension of SORT (Simple Online and Real Time Tracker) based on deep neural network, which is used for efficient real time multi object tracking.

The proposed algorithm is tested using some of the real pedestrian traffic surveillance videos and the performance measure, namely, F1 score, is obtained as 97% for pedestrian detection. Also, the experimental results obtained by using the proposed

algorithm are compared with that obtained by using other multi-object tracking algorithms, which are tested on the 2DMOT2015 and MOT16 datasets.

The organization of paper comprises five sections. In the Sect. 2, a review of the related work is presented. The proposed multi-object tracking algorithm is described in the Sect. 3. In the Sect. 4, the analysis of experimental results is given, while the conclusions are given in the Sect. 5.

## 2 Related Work

The main steps in an object tracking algorithm are: (i) object detection, (ii) multi-target tracking, and (iii) target object counting. Specifically, in the current literature, the pedestrian detection methods are categorized into: (i) object detection methods based on background modeling and (ii) object detection methods based on statistical learning. In a background modeling method, the foreground moving object is extracted, which is then classified as a pedestrian using a classifier, e.g. GMM based algorithm [13] and the ViBe algorithm [14]. Such an algorithm is faster, but it fails to cope with the issues related to variations of illumination and the camera jitters. In addition, it is challenging to discriminate dense objects or occluded objects. On the contrary, an algorithm based on statistical learning can effectively address the issues related to occlusion and changes in illumination and hence, yield higher accuracy, e.g. HOG based algorithm [15]. In recent years, the deep learning technique is used to enhance the pedestrian detection algorithm performance by employing convolutional neural network [16]. Deep learning based detection methods are classified as two-stage methods and one-stage methods. Shaoqing Ren [17] proposed R-CNN and its variants [18, 19], which are the examples of two-stage methods, which effectively locate and detect objects with higher accuracy. The SSD detection algorithm proposed by Wei Liu et al. [20] and variants [21, 22] of this algorithm, are the examples of one-stage algorithms, which achieve better accuracy and attain real-time performance.

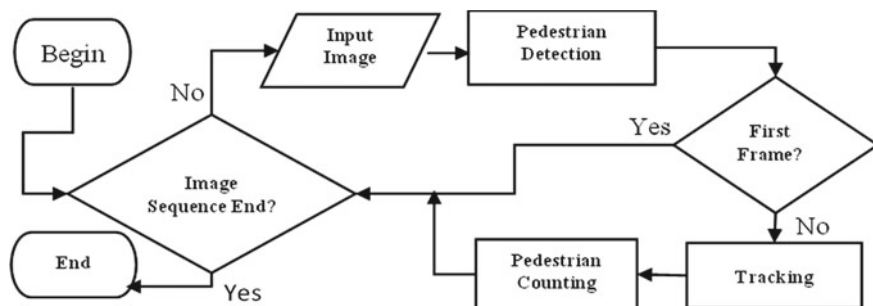
Methods of multi-object tracking are broadly categorized as detection-based-tracking (DBT) and detection-free-tracking (DFT), depending on which method of initialization is used. In DFT, the targets need manual labelling and thereafter tracking is done in succeeding frames [23, 24]. In DBT, tracking by detection is achieved by target detection in each image frame and then by insertion of the detected targets in the tracklets [25–27]. Hence, for pedestrian traffic statistics applications, DBT methods are found to be more suitable, wherein new targets appear and the old ones disappear regularly. Further, depending on the mode of processing, the methods of multi-object tracking can be categorized as online and offline methods. In an online method, the current image frame of the sequence and its some preceding frames are used [28, 29]. Such a method is more suitable for real-time applications. An offline method needs to use some frames succeeding to the current frame in the sequence [30, 31], which makes it apt for post-analysis of a video.

The appearance model, which associates tracklets and detections, is often employed in multi-target tracking, wherein the ID stitches are effectively suppressed. A multi-object tracking method based on an appearance model using HoG descriptor is proposed by Ullah et al. [32]. Bae et al. [33] have proposed a deep learning approach to build an appearance model capable of distinguishing multiple objects with significant appearance variations. Also, the motion model is equally important as much as the appearance model. Often, the object motion in the frame sequence is relatively flat and, hence, the trajectory estimation for the object motion helps to predict the location of the object in the next image frame, which thereby reduces the area of search and also realizes the tracking.

Originally, in the counting method for pedestrian flow, a line is marked on the road and, whenever the mark is covered, it indicates that a pedestrian has passed by. Using this idea, Kryjak et al. [34] proposed a pedestrian counting method using a virtual line. However, in case of targets that hover around the virtual line, it affects the accuracy of the count significantly. Later, Xu et al. [2] proposed a counting method using pairs of virtual lines, wherein two virtual lines are considered and the count of pedestrians crossing over both the virtual lines in the pair is done.

### 3 Methodology

The proposed method consists of three processing steps, namely, pedestrian (object) detection, multi-pedestrian (object) tracking, and pedestrian (object) counting, which are described below. Figure 1 shows the flowchart of the proposed method.



**Fig. 1** Flow chart of the proposed method

### 3.1 Pedestrian Detection

For the pedestrian detection step, YOLOv3 detection framework [35] is modified, wherein a compressed VGG network [36] is used as the front network of the framework, and thereby reduce the network complexity from 65 to 39 Bflops. The main object for detection algorithm is pedestrian. It is achieved by using a clustering algorithm. The usage of very small anchor settings does not affect detection performance and reclustering the size of anchors is accomplished by using K-Means algorithm. The computational complexity is further reduced to 34 Bflops by removing the final FPN structure of the YOLOv3 network because of the targets of larger scale in the front-down camera. To avoid the severe occlusion between the pedestrian targets, soft-NMS method is used. It overcomes the drawback of traditional non-maximum suppression (NMS) [37]. In soft-NMS method, the bounding box, wherein IOU is greater than the threshold and confidence is found to be lower, is not removed directly but its confidence is reduced. Hence, it is more unlikely that the correctly detected targets are removed incorrectly due to the dense targets. Thus, the false positives of detected targets are correctly removed and also the false negatives are reduced.

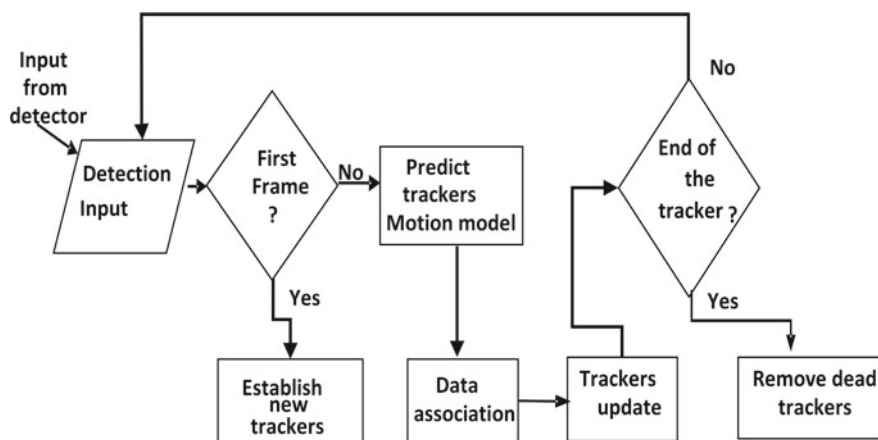
### 3.2 Multi-Pedestrian Tracking

In the pedestrian tracking step, the Deep SORT algorithm is used for tracking multi-targets, which is based on distance, velocity and persons' looks. Deep SORT computes the in-depth features for every bounding box and, using similarities between deep features of targets, the tracking logic is developed. First, a classifier is built for the given dataset, which is trained well to achieve an acceptable degree of accuracy, and then remove the final classification layer. It means that, in view of a classical architecture, this process ends up with a dense layer producing a single feature vector, which needs to be classified. Such a feature vector becomes the "appearance descriptor" of the object. The implementation of radial basis functions as activation functions in the hidden layer achieves higher degree of nonlinear separability leading to higher classification accuracy. Thus, Deep SORT is an elegant and most popular Object Tracker, due to the powerful deep learning technique combined with a simple distance metric. The Fig. 2 illustrate an example of multi-pedestrian tracking.

Association of the bounding boxes with the existing tracks is implemented by using the Hungarian algorithm. The assignment cost matrix is computed as Intersection-Over-Union (IoU) distances to maximize the overlap between predictions and detections. In the first image frame of the sequence, detection results are used to build motion models, which are then used as new tracklets for initializing the sequence. The tracking state is entirely determined by the prediction matrix when unmatched tracklets are found as output, and the predicted tracklets move in the straight line as determined by the state variables. This process mitigates effects of occlusion or detector failure quickly. Failing to match any detected target in multiple



**Fig. 2** Multi-pedestrian tracking example



**Fig. 3** Flow chart of the tracking algorithm

sequential frames implies the disappearance of the target being tracked by the tracklet, and hence, the tracklet is discarded. Figure 3 depicts the flow chart of the tracking algorithm.

### 3.3 Counting of Pedestrian

The proposed pedestrian counting algorithm makes use of virtual blocks. It works in the same manner as the counting method that uses virtual double lines. The virtual block counting algorithm counts the pedestrians based on the sequence of blocks crossed over by the pedestrian. This process of counting reduces the time required for target tracking continuously and achieves robustness with respect to occlusion. The

block-based method is suitable for the counting area with different shapes. When the working application scenarios are taken with front-down cameras, the virtual block method is adequate. With the help of block-based methods, it allows flexibility, like counting part of the road area. One can set the start and end areas for counting, so that count of only the target that enters through the designated entry or leaves from the designated exit is done.

#### 4 Experimental Results and Discussion

For the purpose of experimentation of the proposed algorithm for multi-object tracking for pedestrians, an own video data set is created by capturing videos by front down cameras placed at the entry gate of crowded scenic spots frequented by pedestrians. The dataset contains five 30-min videos of pedestrian traffic scenes with variations in illumination and fog density. The object labelling in the images is done manually. The frame rate of videos is 26 fps. An image is saved after every 15 frames. There are 18,760 annotated images containing a total of 282,680 pedestrian bounding boxes. Out of these images, 2500 images from the two video sequences having different complexities of scenes are used as the experimental test set, while the remaining 16,780 images form the training set. The benchmark datasets, namely, COCO and 2DMOT2015, are also used for the purpose of performance comparison with other methods in the literature. The dataset used for experiment is shown in Table 1.

The performance of the proposed object (pedestrian) detection algorithm is shown in the Table 2. To test the robustness of the detector, the experimentation has been done using the COCO dataset of pedestrian traffic flow and also the own dataset. In comparison with detectors that are trained using only COCO dataset, the robustness of the detector trained with own dataset, which has variations in illumination and fog concentration, is much improved due to enhancement of the recall rate.

The sample results of multiobject tracking algorithm in five different video frame sequences, namely, Seq1, Seq2, Seq3, Seq4, Seq5, are shown in the Fig. 4, which show the effectiveness of the proposed algorithm. In Seq1, pedestrians near the camera and also for away from the camera are detected. In Seq2, pedestrian’s density is high and illumination is moderate. In Seq3, pedestrians under the bright light as

**Table 1** Datasets used for experiments

Dataset	Number of frames	Remarks
COCO	250,000 images	Used for training detector
Own dataset	18,760 images	Used for training detector
2DmOT2015	4400 images	Used for detecting and tracking

**Table 2** Performance of the detector which is trained with two different data sets

Performance measure	Dataset	
	COCO	COCO + own data
Precision	88%	87%
Recall	69%	81%
F1-Score	77%	76%
True positive (TP)	18,796	21,156
False positive (FP)	2544	3300
False negative (FN)	8400	5000
Mean average precision (MAP)	73.88%	84.39%



**Fig. 4** Detector and tracker results for 5 video frame sequences Seq1 to Seq5

well as those under shadow are detected. In Seq4, there is poor illumination. In Seq5, there is bright illumination and cluttered background.

Finally, the performance evaluation of the counting algorithm is carried out in terms of metrics, namely, the precision, recall, and F1-Score, as compared to the ground truth and is presented in the Table 3.



**Table 3** The performance of the counting algorithm in five test videos Seq1 to Seq5 (Fig. 4)

Sequence	Seq1	Seq2	Seq3	Seq4	Seq5	Total
Count Result	570	414	574	950	685	3193
Ground Truth(GT)	577	415	575	965	695	3227
FP	18	14	18	14	12	76
FN	27	14	10	35	30	116
TP	550	399	565	930	665	3109
Precision %	98.21	96.61	96.91	98.51	98.22	97.69
Recall %	95.32	96.61	98.26	96.37	95.68	96.64
F1-Score %	96.74	96.61	97.58	97.42	96.93	<b>97.05</b>

It is observed from the Table 3 that the mean F1-score of the proposed counting algorithm is 97% (approx.), which demonstrates the effectiveness of the proposed counting algorithm.

The performance of pedestrian detector and tracker algorithms has a significant influence on the pedestrian counting results. Therefore, the performance comparison of the proposed method with other methods in the literature has been done choosing the benchmark dataset 2DMOT2015 [38], which is popularly used for the performance evaluation of multi-pedestrian tracking methods. This dataset comprises 22 image sequences, out of which 11 are used for training and remaining 11 for testing. Image sequences come from several popular datasets used for pedestrian detector and tracker, namely, ADL, KITTI, PETS, ETH and TUD.

The indicators often used by researchers for performance evaluation are given in the Table 4 [42]. To verify the overall performance of the proposed detector and tracker, it is compared with other algorithms in the literature, namely, MDP\_Subcnn [39], DMT [40], PFS [42] and Sort [41], in terms of the performance indicators mentioned in the Table 4.

The specific comparison results are shown in the Table 5, which demonstrate the effectiveness of the proposed algorithm.

**Table 4** Indicators used for performance evaluation

Measure	Better	Description
MOTA	Higher	Multiple object tracking accuracy
MOTP	Higher	Multiple object tracking precision
IDF1	Higher	The ratio of correctly identified detections
MT	Higher	Mostly tracked targets
ML	Lower	Mostly lost targets
FP	Lower	The total number of false positives
FN	Lower	The total number of false negatives
IDS <sub>w</sub>	Lower	The total number of identity switches



**Table 5** The performance comparison of the proposed algorithm with other algorithms, namely, MDP\_Subcnn [39], DMT [40], PFS [42] and Sort [41], using 2DMOT2015 dataset

Measure	Algorithm				
	Proposed algorithm	MDP_Subcnn	DMT	Sort	PFS
MOTA	49.5	47.5	44.5	33.4	48.1
IDF1	56.8	55.7	49.2	40.4	45
MOTP	76.2	74.2	72.9	72.1	74.7
MT	0.358	0.3	0.347	0.117	0.327
ML	0.1	0.186	0.221	0.309	0.15
FP	7300	8631	8088	7318	8422
FN	21,950	22,969	25,335	32,615	22,074
IDS <sub>w</sub>	620	628	684	1001	1372

## 5 Conclusion

In this paper, a novel real-time multi-pedestrian tracking is developed, which comprises revamped YOLOv3 for object detection and Deep SORT for object tracking by using movement representation and data association algorithms. The virtual blocks counting method is implemented for efficient multi-object tracking and counting of the pedestrian, in which the problem of occlusion is resolved effectively. The experimentation of the proposed algorithm is carried out on different benchmark datasets and also own dataset of videos of real scenes. The results of experimentation demonstrate that the proposed algorithm yields higher accuracy and has effective real-time performance as compared to other object tracking techniques in the literature and, hence, is a reliable object-tracking algorithm. In future work, the categorization of pedestrians based on age and gender will be considered.

**Acknowledgements** The authors are great full to reviewers for their help full comments and suggestions, which improved the quality of the paper considerably.

## References

1. Del Pizzo L, Foggia P, Greco A, Percannella G, Vento M (2016) Counting people by RGB or depth overhead cameras. *Pattern Recog Lett* 81:41–50
2. Coskun A, Kara A, Parlaktuna M, Ozkan M, Parlaktuna O (2015) People counting system by using kinect sensor. *Proceedings of the IEEE international symposium on innovations in intelligent systems and applications (INISTA)*, Taipei, Taiwan, 2–4 September, pp 1–7
3. Verma NK, Dev R, Maurya S, Dhar NK, Agrawal P (2019) People counting with overhead camera using fuzzy-based detector. In: *Computational intelligence: theories, applications and future directions*. vol 1. Springer: Berlin, Germany, pp 589–601
4. Kopaczewski K, Szczodrak M, Czyzewski A, Krawczyk H (2015) A method for counting people attending large public events. *Multimed Tools Appl* 74:4289–4301

5. Beymer D (2000) Person counting using stereo. In: Proceedings of the workshop on human motion, Austin, TX, USA, 7–8 December, pp 127–133
6. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. In: Proceedings IEEE, pp 127–133
7. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Proceedings of the advances in neural information processing systems, Lake Tahoe, NV, USA, 3–6 December, pp 1097–1105
8. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv 2014, [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
9. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, MA, USA, 7–12 June, pp 1–9
10. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 27–30 June, pp 770–778
11. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA, 21–26 July, pp 2261–2269
12. Bochinski E, Eiselein V, Sikora T (2017) High-speed tracking-by-detection without using image information. In: Proceedings of the 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS), Lecce, Italy, 29 August–1 September, pp 1–6
13. Stauffer C, Grimson WEL (1999) Adaptive background mixture models for real-time tracking. In: Proceedings of the 1999 IEEE computer society conference on computer vision and pattern recognition (Cat No PR00149), Fort Collins, CO, USA, 23–25 June, p 2246
14. Barnich O, Van Droogenbroeck M (2009) A powerful random technique to estimate the background in video sequences. In: Proceedings of the IEEE international conference on acoustics, speech and signal in processing ICASSP 2009, Taipei, Taiwan, 19–24 April, pp 945–948
15. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition CVPR 2005, San Diego, CA, USA, 20–25, vol 1, June, pp 886–893
16. Zhou C, Yuan J (2018) Bi-box regression for pedestrian detection and occlusion estimation. In: Proceedings of the european conference on computer vision (ECCV), Munich, Germany, 8–14 September, pp 135–151
17. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 1137–1149
18. Dai J, Li Y, He K, Sun J (2016) R-fcn: object detection via region-based fully convolutional networks. In: Proceedings of the advances in neural information processing systems, Barcelona, Spain, 5–10 December, pp 379–387
19. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: Proceedings of the CVPR, Honolulu, HI, USA, 21–26 July, vol 1, p 4
20. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) Ssd: single shot multibox detector. In: Proceedings of the european conference on computer vision, Amsterdam, The Netherlands, 8–16 October 2016, Springer: Berlin, Germany, pp 21–37
21. Fu CY, Liu W, Ranga A, Tyagi A, Berg AC (2017) DSSD: deconvolutional single shot detector. arXiv 2017. [arXiv:1701.06659](https://arxiv.org/abs/1701.06659)
22. Zuoxin Li, Zhou F (2017) FSSD: feature fusion single shot multibox detector. arXiv 2017. [arXiv:1712.00960](https://arxiv.org/abs/1712.00960)
23. Hu W, Li X, Luo W, Zhang X, Maybank S, Zhang Z (2012) Single and multiple object tracking using log-Euclidean Riemannian subspace and block-division appearance model. IEEE Trans Pattern Anal Mach Intell 34:2420–2440
24. Van der Maaten Z (2013) Structure preserving object tracking. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Portland, OR, USA, 23–28 June, pp 1838–1845

25. Berclaz J, Fleuret F, Turetken E, Fua Berclaz P (2011) Multiple object tracking using k-shortest paths optimization. *IEEE Trans Pattern Anal Mach Intell* 33:1806–1819
26. Breitenstein MD, Reichlin F, Leibe B, Koller-Meier E, Van Gool L (2009) Robust tracking-by-detection using a detector confidence particle filter. In: *Proceedings of the 2009 IEEE 12th international conference on computer vision*, Kyoto, Japan, 29 September–2 October, pp 1515–1522
27. Ess A, Leibe B, Schindler K, Van Gool L (2009) Robust multiperson tracking from a mobile platform. *IEEE Trans Pattern Anal Mach Intell* 31:1831–1846
28. Choi W, Pantofaru C, Savarese S (2013) A general framework for tracking multiple people from a moving camera. *IEEE Trans Pattern Anal Mach Intell* 35:1577–1591
29. Khan Z, Balch T, Dellaert F (2005) MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Trans Pattern Anal Mach Intell* 27:1805–1819
30. Kuo CH, Huang C, Nevatia R (2010) Multi-target tracking by online learned discriminative appearance models. In: *Proceedings of the 2010 IEEE conference on computer vision and pattern recognition (CVPR)*, San Francisco, CA, USA, 13–18 June 2010, pp 685–692
31. Milan A, Roth S, Schindler K (2014) Continuous energy minimization for multi-target tracking. *IEEE Trans Pattern Anal Mach Intell* 36:58–72
32. Ullah M, Cheikh FA, Imran AS (2016) Hog based real-time multi-target tracking in the bayesian framework. In: *Proceedings of the 2016 13th IEEE international conference on advanced video and signal based surveillance (AVSS)*, Colorado Springs, CO, USA, 23–26 August, pp 416–422
33. Bae SH, Yoon KJ (2018) Confidence-based data association and discriminative deep appearance learning for robust online multi-object tracking. *IEEE Trans Pattern Anal Mach Intell* 40:595–610
34. Kryjak T, Komorkiewicz M, Gorgon M (2014) Hardware-software implementation of vehicle detection and counting using virtual detection lines. In: *Proceedings of the 2014 conference on IEEE design and architectures for signal and image processing (DASIP)*, Madrid, Spain, 8–10 October 2014, pp 1–8
35. Redmon J, Farhadi A (2018) Yolov3: an incremental improvement. *arXiv* 2018, [arXiv:1804.02767](https://arxiv.org/abs/1804.02767)
36. He Y, Zhang X, Sun J (2017) Channel pruning for accelerating very deep neural networks. In: *Proceedings of the international conference on computer vision (ICCV)*, Venice, Italy, October, vol 2, pp 22–29
37. Bodla N, Singh B, Chellappa R, Davis LS (2017) Soft-nms—improving object detection with one line of code. In: *Proceedings of the 2017 IEEE international conference on computer vision (ICCV)*, Venice, Italy, 22–29 October, pp 5562–5570
38. Leal-Taixé L, Milan A, Reid I, Roth S, Schindler K (2015) Motchallenge 2015: towards a benchmark for multi-target tracking. *arXiv* 2015, [arXiv:1504.01942](https://arxiv.org/abs/1504.01942)
39. Xiang Y, Alahi A, Savarese S (2015) Learning to track: online multi-object tracking by decision making. In: *Proceedings of the IEEE international conference on computer vision*, Región Metropolitana, Chile, 11–18 December, pp 4705–4713
40. Kim HU, Kim CS (2016) CDT: cooperative detection and tracking for tracing multiple objects in video sequences. In: *Proceedings of the european conference on computer vision*, Amsterdam, The Netherlands, 8–16 October 2016, Springer: Berlin, Germany, pp 851–867
41. Bewley A, Ge Z, Ott L, Ramos F, Upcroft B (2016) Simple online and real-time tracking. In: *Proceedings of the 2016 IEEE international conference on IEEE image processing (ICIP)*, Phoenix, AZ, USA, 25–28 September, pp 3464–3468
42. He M, Hui B (2019) Pedestrian flow tracking and statistics of monocular camera based on convolutional neural network and kalman filter. In *Appl Sci MDPI* 1624:9

# Student Engagement Awareness Dashboard for Asynchronous E-learning Environment



Abdalganiy Kebede and Samit Bhattacharya

**Abstract** In asynchronous e-learning, students find themselves all alone in their learning process. This feeling of loneliness mostly results in frustration and loss of motivation, leading to a high drop-out rate. Tracking engagement while the student is learning allows intervention at an appropriate time. However, instructors may get overwhelmed by data provided to them in the online courses. In our work, we propose a student engagement visualization dashboard that visualizes the instantaneous engagement levels every minute, visualizes trends of student engagement levels and filters and displays the least engaged learner to address the above challenges. We visualized engagement levels that were predicted from a model. We also evaluated the usability of this visualizer in a controlled experiment and found out that the perceived usefulness by the teachers was high. The visualizer allows teachers to gain insight into the engagement levels of all the students at a glance. It also allows the teacher to take immediate action.

**Keywords** Asynchronous e-learning · Engagement prediction · Visualization · HCI

## 1 Introduction

E-learning has been defined as a platform or mode of learning allowing learners to attend a course online where they will neither meet the instructor face to face nor access on-campus learning materials [32]. E-learning provides an important means of education which can reach masses irrespective of their locations all over the world [1]. The face to face learning mode is known to be traditional in contrast to the online learning which makes use of modern information technologies [18].

Online learning can be synchronous or asynchronous. Asynchronous online communication occurs in a time-independent environment, whereas synchronous

---

A. Kebede (✉) · S. Bhattacharya  
Indian Institute of Technology, Guwahati, Assam, India

S. Bhattacharya  
e-mail: [samit@iitg.ac.in](mailto:samit@iitg.ac.in)

conferencing systems are time-dependent systems. Learners in the synchronous system faced challenges such as time constraints [20]. Asynchronous online learning allows learning independent of time, place and pace [18]. One of the disadvantages of asynchronous e-learning is students find themselves all alone in their learning process. This feeling of loneliness results in frustration and loss of motivation, leading to a high drop-out rate [15]. More over, the absence of human supervision is always a concern as a student cannot be monitored for losing interest or not getting engaged in the e-learning session [1].

E-Learning systems could be improved by tracking students' engagement that in turn would allow personalized interventions at appropriate times to reengage students [10]. Student engagement is about students putting time, energy, thought, and effort and to some extent feelings into their learning [12]. Student engagement in online courses is generally poorer than in face-to-face courses [27]. Poor student engagement poses risks to both the students themselves and their instructor [27]. Engagement is the key factor for successful learning [21]. Learner's engagement can be traced from the interaction of the learner with online content through a learning tool [4]. The utilization of facial features can also be used to automatically detect student engagement [6].

Engagement prediction models, coupled with a student engagement visualization system, can help instructors evaluate student engagement to provide additional interventions for students [13]. Tracking engagement while the student is learning allows intervention at an appropriate time [10]. Visualizing data helps stakeholders easily see trends to understand what has been happening [5].

Instructors are overwhelmed by data reports provided to them in online courses [5]. Such challenges were solved in the literature by classifying the student into different classes of engagement levels using the model they built. Coffin et al. [11] applied a model and classified students into three categories: auditors, active and qualified and visualized the outputs of the model predictions. However, they did not consider further classifying or filtering students in one of these categories. After classifying a learner as auditors, there may be a large number of students in this particular category. Then, the task of identifying the least engaged learner from such larger numbers of students in this category is another challenge. We have not seen any work that reported such work.

In our work, we propose a student engagement visualization dashboard that visualizes the instantaneous engagement levels every minute, visualizes trends of student engagement levels and filters and displays the least engaged learner to address the above challenges. The dashboard is based on a student engagement prediction model, which we also developed. We also performed the validation of these proposed visualizers in controlled experiment.

The paper is organized as follows. A literature review is presented in Sect. 2. The model for student engagement prediction that we built is described in Sect. 3. The proposed visualizer is reported in Sect. 4 with the validation of the visualizer followed by a discussion on the pros and cons of the proposed work. Section 5 concludes the paper.

## 2 Literature Review

We reviewed papers concerning two issues: i. visualizing student engagement levels based on prediction model or algorithm ii. visualizing student engagement level without using prediction models.

Liu et al. [16] estimated student engagement through two algorithms they developed before visualizing it. The algorithms are point-based algorithm (pbA) and intensity-based algorithms (ibA). They implemented three types of visualizations: point-based visualization, line-based visualization and height-based visualization. Liu et al. [16] used a different color on height-based visualization to represent different tasks. They described a learning analytic system called Tracer, which derives behavioral engagement measures and creates visualizations of behavioral patterns of students writing on a cloud-based application and a novel learning analytic (LA) system that collects behavioral data of users writing, estimates the level of engagement. In Liu et al. [16], the minimum time they used for the interaction before visualizing engagement was two minutes. Hussain et al. [13] visualized low and high engagement levels after predicting the engagement levels using a model. They developed a dashboard to visualize the number of activities in VLE, individual student engagement in each assessment and the percentage of low and high-level engagement using a chart [13]. Coffrin et al. [11] clustered students into subpopulation to gain additional insight. The subpopulations or the groups were: auditors, active and qualified students based on learner grade performance within the first weeks of the course. They were able to reveal a more detailed story of student engagement after they divided the students' related data into the groups. They implemented a histogram to visualize weekly student participation. They also applied a cumulative distribution plot which does not require a fixed number of bars, unlike histograms.

Carrillo et al. [9] visualized only indicators of engagement but not measured engagement from a prediction model. They used the number of logins on a learning application by time period, the number of times a learning resource was accessed, and the time spent on a learning document for behavioral indicators and cognitive indicators such as which node(s) or link(s) did s/he delete? When did s/he do it? How did s/he modify the structure of the mind map document? Pesare et al. [21] did not apply student engagement prediction model to predict the engagement before visualizing it. They implemented scatterplot visualization to display trends and relationships in a cloud of points and a linear visualization, a time series to visualize details of a particular student in terms of trends and distribution. Moreover, a pie chart to display the distribution of interaction. In our work, we build on the visualizations from the works of Liu et al. [16] and Pesare et al. [21].

In the earlier works by Hussain et al. [13], two levels of engagement were used. Coffin et al. [11] clustered students into three categories. However, they did not consider further classifying or filtering students in these categories. Considering that each category may contain large number of students, we believe a larger number of category would be beneficial. As a result, we propose a prediction model that categorizes students into either of four categories. Another issue is the frequency

of updating of the visualizer. Liu et al. [16] used a 2-min interval for updating. In our case, the visualizer that we built visualizes engagement levels every minute, providing the instructor a better-view of the student status. Moreover, we propose to visualize the least engaged learner after classifying students into four classes of engagement levels. This is so since there may be a large number of students in a class and automatic identification of the least engaged learner may help the instructor save time and effort.

### 3 Proposed Student Engagement Prediction Model

In order to build the visualizer, we developed a student engagement prediction model. The first step to build the engagement model was to define engagement levels, which is presented next.

#### 3.1 Engagement Levels

We propose four levels of engagement, based on Whitehill et al. [31] and Kaur et al. [14]: Very High (VH), High (H), Low (L) and Very Low (VL). Earlier works suggested that engagement levels can be predicted reliably from 10-s video clips [31]. When the student seems uninterested and looks away from the screen frequently, we classify that engagement as VERY LOW (VL). When the student is clearly not into the task and moves restlessly in the chair, we classify that engagement as LOW (L). When the student seems to like the content, and requires no admonition to stay in the task, we classify that engagement as HIGH (H). When the student stared at the screen and was focused, and could be highly praised for his/her level of engagement in the task, we classify that engagement as VERY HIGH (VH).

#### 3.2 Details of Empirical Study

We built the proposed predictive model from empirical data. In our study, Moodle was chosen as a learning management system to allow students to interact with learning activities. We used version 3.5 on Ubuntu 18.04.

For every task and participant, the log files of the interaction with the LMS are recorded automatically, at 5 min interval.

In our study, we used a facial emotion recognition tool called *clmtrackr* to capture the facial emotion automatically [19, 24]. We accessed the source code from web URL: <https://github.com/auduno/clmtrackr>. We installed version v1.1.2. The tool calculates the recognition rate of six basic emotions: disgust, angry, fear, sad, happy and surprise by getting the model fitting score of the classified image as an emotion.

It produces downloadable log file as comma separated value (csv) at the end of the session and offers dialog box in the browser window. It uses timer of elapsed time.

Four tasks were designed and implemented, which were content viewing, quiz, assignment and discussion forum as online tasks, using a course on Descriptive Statistics with worked-out examples. Content viewing, quiz, and assignment were designed to detect behavioral engagement as explained by Wang and Degol [29]. According to Husain et al. [13], content viewing, discussion forum and quiz are significantly correlated with engagement. Assignment is the most used indicator of engagement [2]. We have also applied discussion forum to be used as one of the tasks to detect collaborative engagement [22]. According to Redmond et al. [23], individual interaction of learners with each other has been the main influencer of engagement. Moreover, educational technologies such as discussion boards could enhance the learning experience because discussions are captured and can be reviewed later by students and instructors [25]. Summary of the tasks performed is presented in Table 1.

We used 12 graduate students as participants for the study. All were volunteers. Their average age was 33.6 years, with minimum 24 years and maximum 39 years. There were 10 males and 2 females. The participants interacted with Moodle in two sessions, minimum of 15 min each. Three of them had prior experiences with the LMSs.

Participants were briefed about the tasks. Afterwards, they performed the tasks mentioned in Table 1. Each participant performed the tasks over two sessions of 15 min each. The log file of the interaction with Moodle was later extracted at the end of the two sessions. This log file was divided in 5 min sample. At the same time, the log file of the rate of facial emotion was downloaded every 5 min. Each interaction of the participants in the experiment was recorded with a screen recording software. The recorded video interaction was used for labeling the engagement levels of each participant for analysis purpose. The labelling task was performed by viewing 10-s video clips from the recorded interaction videos and assigning a number (between 1 and 4) to rate each video frame, as explained in Table 2.

**Table 1** Summary of tasks performed

Task	Type of feature	Specific features captured
Lesson, quiz and assignment	Behavioural features	Number of content view (NCV), time to read content (TRC), score (SC) and time to submit assignment (TA)
Discussion forum	Collaboration feature	time between posts and replies (TPR), time in the forum (TF) and number of replies (NR)
A student sits in front of a computer with a webcam and allowing face tracking while performing the above tasks	Emotional feature	Rates of the basic emotions: disgust (DIS), angry (ANG), fear (FEA), sad (SAD), surprise (SUR) and happy (HAP)



**Table 2** Empirical video data rating scheme used in the study

Engagement intensity	Meaning	Point given
Very low (VL)	Not engaged at all	1
Low (L)	Barely engaged	2
High (H)	Engaged in the content	3
Very high (VH)	Very engaged	4

### 3.3 Proposed Predictive Model for Engagement

We used Pearson correlation coefficient to analyze data. We assumed values between 0.3 and 0.7 (-0.3 and -0.7) indicate a moderate positive (negative) linear relationship. Values between 0.7 and 1.0 (-0.7 and -1.0) indicate a strong positive (negative) linear relationship [7]. We used MS Excel 2010 for the analysis.

The correlation was computed between very low (VL) engagement (ENG-VL) levels, low (L) engagement levels (ENG-L), high (H) engagement levels (ENG-H) and very high (VH) engagement levels (ENG-VH) and the features. Table 3 summarizes the significant features which were identified from the correlation analysis result.

The student engagement prediction model that we built was based on the significant features identified in Table 3. Using non-linear regression analysis, we obtained a set of 4 equations, one each for a particular engagement class relating to the significant features, as our proposed model. Based on the model, we computed the value

**Table 3** Summary of the correlation analysis results

Engagement level	Type of feature	Significant feature (see Table 1)	Number of significant features
Very low (VL)	Collaboration		2
	Behavioural	TA	
	Emotional	SUR	
Low (L)	Collaboration	TPR	5
	Behavioural	TA, TRC, SC	
	Emotional	SUR	
High (H)	Collaboration	NR, TPR, TF	10
	Behavioural	TA, TRC, NCV	
	Emotional	ANG, SAD, SUR, HAP	
Very high (VH)	Collaboration	NR, TPR	9
	Behavioural	TA, TRC, NCV, SC	
	Emotional	ANG, SUR, HAP	

**Table 4** The proposed model and the engagement level ranges

Proposed Model (relationships between engagement class and significant features)	Range	
	Min	Max
$V L = 0.02 \times SUR + 0.23$ where $2.9 \leq SUR \leq 11$	= 0	= 2
$L = 0.06 \times SC^2 + 0.05 \times SC + 1.7 \times TRC^2$ $- 7.5 \times TRC + 0.06 \times SUR + 10,$ Where, $TRC \geq 2.4$ and $2.9 \leq SUR \leq 11$	>2	= 7
$H = 15 \times NR^2 - 15 \times NR + 2 \times TF^2$ $- 3 \times TF - 0.02 \times NCV^2 + 2 \times NCV - 1.28 \times TRC^2$ $+ 9.5 \times TRC - 0.03 \times ANG + 0.01 \times SUR - 0.123$ $\times HAP + 0.43 \times SAD + 35.5,$ Where $0 \leq NR \leq 0.4, 0.4 \leq TF \leq 0.8,$ and $0.13 \leq SAD \leq 21$	>7	= 62
$VH = 15.3 \times NR + 0.4 \times SC^2 - 4 \times SC + 0.02 \times NCV^2$ $- 2 \times NCV - 0.8 \times TRC^2 - 0.7 \times TRC$ $+ 0.02 \times ANG - 0.09 \times SUR + 0.09 \times HAP + 232,$ Where, $3 \leq NCV \leq 9.2, 1.8 < TRC \leq 3.4,$ and $0 \leq HAP \leq 43$	>62	= 254

ranges for each of the four classes of engagements. The equations and the engagement level ranges are shown in Table 4.

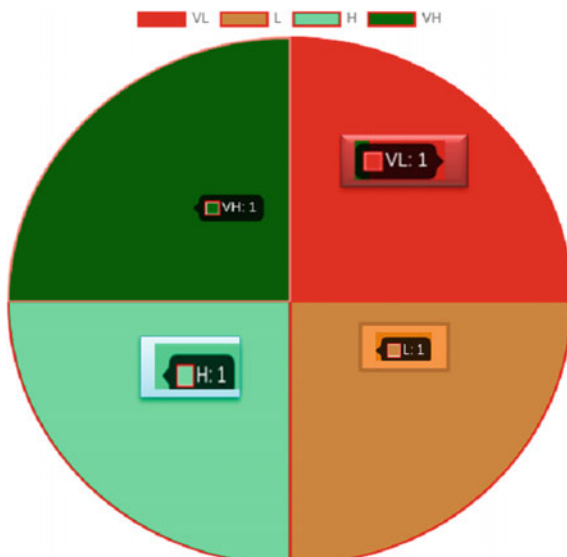
4 The Proposed Visualizer

We visualized student engagement levels after predicting it using the model we built. We built three types of visualizations: *instantaneous engagement level visualization*, *engagement level trend visualization* and *least engaged learner visualization*. In the subsequent section, we explain each of these visualizations.

4.1 Instantaneous Engagement Level Visualization

Instantaneous engagement level visualization refers to displaying the number of students with a particular engagement level every minute. We used a pie chart to visualize the instantaneous engagement levels of the students [26]. There are four sectors of the pie chart as there are four levels of engagement, rendered with four different colors. The variation of the colors was according to the intensity of engagement levels: deep green for VH, light green for H, light red for L and dark red for VL.

**Fig. 1** Pie chart visualization of instantaneous engagement level exhibited by four students



The size of the sector is proportional to the number of students at the corresponding engagement level. An example is shown in Fig. 1.

## 4.2 Engagement Level Trend Visualization

Engagement level trend visualization refers to displaying the number of students experiencing the four engagement levels for the last 30 min. We used a line graph to visualize the number of students with the engagement levels, as line graphs are useful to support awareness of progress during specified periods of time [26]. We plotted the number of students with the engagement levels in the y-axis and the time of visualization in a minute on the x-axis. Figure 2 depicts an example of the trend visualization.

## 4.3 Least Engaged Learner Visualization

Least engaged learner visualization refers to displaying through filtering only those students who are critically disengaged for the last 30 min. The visualization was done using textual format [17], which is a list. In order to obtain the list, the instructor can click on any instance on the trend line view. The list is in the order of decreasing disengagement level. The top ones are the most disengaged learners who need immediate intervention. We counted the number of the very low (VL) and low (L) engagement

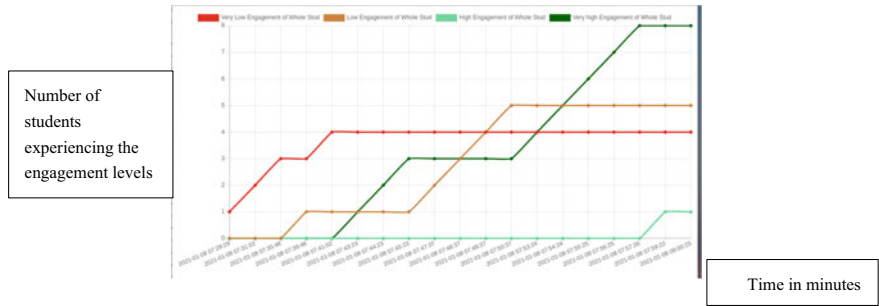


Fig. 2 An example of trend visualization using line graph

Userid	Firstname	Lastname	Engagement Level Predicted	Number of Low or Very low engagement levels(States)(Frequency)
21	MOHAMMAD	ABDULAUADIR	L	7
13	Sibu	Sibu	VL	5
7	Sayyed	Nayyabrasul	L	4
22	Tufa	Feyisa	L	4
17	Abdulwahid	Abdulwahid	VL	3
14	Abdisa	Abdisa	VL	2
20	Masresha	Masresha	L	2
23	Mabratu	Mabratu	VL	1

Fig. 3 Visualizing least engaged learners

states of each of the learners whose engagement levels were found in these two levels for the previous 30 min. We then sorted the frequency of the engagement states. The one with the maximum was assumed to be the least engaged learner. An example of the visualization is shown in Fig. 3.

4.4 Implementation Details

The student engagement visualization was developed using chart.js. It is an open-source, tiny, fast, easy to use, library supporting six chart types: doughnut, pie, polar, line, bar and radar. Chart.js uses HTML5 Canvas [8]. We accessed it from <https://github.com/chartjs/Chart.js/releases/tag/v2.9.4> and installed the version: v2.9.4. A simplified API is one benefit to the canvas-based chart.js [28].

4.5 Empirical Validation and Discussion

The visualizer was supposed to be used by the instructors in real-time. Therefore, it was important to ensure that the interface is usable. In order to measure the usability of

the visualizer, we performed an empirical validation study with the System Usability Scale (SUS) [3] instrument.

There were 10 university students who participated in this study as volunteers. All were males. The participants' age ranged from 24 to 42 years, with an average age of 32 years. All of them had teaching experience, as teaching assistants or as teachers. None had any prior experience of using such visualizer. However, five of them were familiar with LMSs. It may be noted that the participants were different from the participants, who took part in the earlier study, described in Sec 3.2.

We created a simulated online learning setting to perform the study. In the setting, one participant acted as a "teacher" and the remaining nine participants were the "students". We did it for ten times, so that each participant can use the visualizer and provide the ratings. Each "teacher" took fifteen minutes in front of the screen, interacting with the three visualizations. The role of the "students" was to interact with Moodle, as in the previous study described in Sec 3.2. During such interaction, the face tracking tool was also used as before. In that way, we predicted the student engagement levels in real-time, using our proposed model, and rendered those on the teacher's screen for visualization. At the end of the fifteen minutes interaction, the teacher provided ratings for the SUS items.

We analyzed the ratings and converted to SUS score for each participant [3]. The minimum SUS score was 75, with an average score of 89.5. The score indicates that the usability of the visualizer is high.

## 4.6 Discussion

In our study, we visualized engagement levels that were predicted from a model. The model predicted a student engagement into one of the four engagement levels every minute. One of the benefits of the visualizer is that it worked seamlessly with an existing and popularly used LMS, Moodle. This is a benefit because people prefer to use tools they know [30].

There were three types of visualizations implemented. The first was instantaneous student engagement visualization. The strategy used to render this level, including the color scheme and the dynamic sector size, were in accordance with the general expectations of a user, making them natural and intuitive. Similar strategies we adopted in the second visualizer as well, in which the temporal trends are displayed using line graphs. We used an interval of 30 min for updating of the trend, which we believe is an optimal interval, as per our empirical observations during the studies. Also, the use of the line graphs for such visualizations is in accordance with the general practice reported in the literature. Overall, we believe that these contribute towards making the visualizer "easy to learn", "easy to understand" and "easy to remember". Moreover, to reduce the information burden on the instructor to determine the students who need further monitoring, we proposed to visualize the least engaged learner. We have not seen any work that reported such finding. The interval of 30 min that we have chosen is based on our empirical experience. The use of list for such visualization

is also in accordance with general findings reported in the literature. Accordingly, we believed the third visualizer is also very useful and usable to an instructor to determine and predict the potential drop-out cases early for further intervention. The high SUS score we observed in our study indicates that all the design decisions that we have made led to a “usable” product.

Although the proposed visualizer was found to be usable, there are few points of concerns as well. First of all, the model-building and evaluation of the visualizer were done with a small group of users. More detailed studies with larger empirical data may be requirement for refinement of the model as well as conclusive validation of the visualizer. Another limitation of this study is that the effect of the visualizer on learning outcomes was not evaluated, which may prove very much useful to convince the instructors for its widespread use. All these issue we would like to take up for further studies in future.

## 5 Conclusion

In this paper, we proposed a model to predict student engagement levels in asynchronous e-learning environment. We also proposed a dashboard for the instructors to visualize the predicted states. There were three types of visualizations implemented. The first was instantaneous student engagement visualization, which allows for visualization of the number of students experiencing each of the four engagement levels every minute. The second visualization was engagement levels trend visualization, using line graphs. Finally, it also supports visualization of the least engaged learners. We evaluated the usability of this visualizer and found out that the perceived usefulness by the teachers was high, as revealed by the SUS score. In future, we plan to perform more experiments to collect more empirical data for further refinement and validation of both the model and the visualizer. Also, we would like to work on testing the visualizer in real-settings and estimate the effect it has on the overall learning outcome.

## References

1. Al-Alwani A (2016) A combined approach to improve supervised e-learning using multi-sensor student engagement analysis, *Amer J Appl Sci*
2. Motz B, Quick J, Schroeder N, Zook J, Gunkel M (2019) The validity and utility of activity logs as a measure of student engagement. In: *The 9th International learning analytics and knowledge conference (LAK19)*, March, 2019, Tempe, AZ, USA, pp 10. <https://doi.org/10.1145/3303772.3303789>
3. Barnum CM (2011) *Usability testing essentials: ready, set–test*/Carol Barnum. Morgan Kaufmann Publishers, Burlington, MA
4. Beck JE (2004, August). Using response times to model student disengagement. In: *Proceedings of the ITS2004 workshop on social and emotional intelligence in learning environments*, vol. 13. pp 20

5. Bodily R, Graham CR, Bush MD (2017) Online learner engagement: opportunities and challenges with using data analytics. *Edu Technol* 10–18
6. Bosch N (2016, July) Detecting student engagement: human versus machine. In: Proceedings of the 2016 conference on user modeling adaptation and personalization, ACM, pp 317–320
7. Ratner B (2004) Statistical modeling and analysis for database marketing. Effect Techniq Mining Big Data-Chapman and Hall\_CRC
8. Caldarola EG, Rinaldi AM (2017) Big data visualization tools: a survey. Research Gate
9. Carrillo R, Lavoué É, Prié Y (2016, April) Towards qualitative insights for visualizing student engagement in web-based learning environments. In: Proceedings of the 25th international conference companion on world wide web, pp 893–898
10. Cocea M, Weibelzahl S (2011) Disengagement detection in online learning: validation studies and perspectives. *IEEE Trans Learn Technol* 4(2):114–124
11. Coffrin C, Corrin L, De Barba P, Kennedy G (2014) March. Visualizing patterns of student engagement and performance in MOOCs. In : Proceedings of the fourth international conference on learning analytics and knowledge, pp 83–92
12. Dixon MD (2015) Measuring student engagement in the online course: The Online Student Engagement scale (OSE). *Online Learn* 19(4):n4
13. Hussain M, Zhu W, Zhang W, Abidi SMR (2018) Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Comput Neurosci*
14. Kaur A, Mustafa A, Mehta L, Dhall A (2018, December) Prediction and localization of student engagement in the wild. In: 2018 Digital image computing: techniques and applications (DICTA). IEEE, pp 1–8
15. Leo et al (2009) Online synchronous instruction: challenges and solutions. *Electron J Electron Learn*
16. Liu M, Calvo RA, Pardo A, Martin A (2014) Measuring and visualizing students' behavioral engagement in writing activities. *IEEE Trans Learn Technol* 8(2):215–224
17. Mazza R, Bettoni M, Faré M, Mazzola L (2012) Moclog—monitoring online courses with log data
18. Nortvig AM, Petersen AK, Balle SH (2018) A literature review of the factors influencing e-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electron J e-Learn* 16(1):46–55
19. Khazan O (2014) This app reads your emotions on your face. The Atlantic. <https://www.theatlantic.com/technology/archive/2014/01/this-app-reads-your-emotions-on-your-face/282993/> Accessed 25 November 2019
20. Park , Bonk (2007) Synchronous learning experiences: distance and residential learners' perspectives in a blended graduate course. *J Interact Online Learning*
21. Pesare E, Roselli T, Rossano V (2016) Visualizing student engagement in e-learning environment. In: 22th International Conference on Distributed Multimedia Systems (DMS), pp 26–33
22. Rabbany R, Elatia S, Takaffoli M, Zaïane OR (2014) Collaborative learning of students in online discussion forums: a social network analysis perspective. In *L Educational data mining*, Springer, Cham, pp 441–466
23. Redmond et al (2018) An online engagement framework for higher education, *Online Learn J* 22(1)
24. Robal T, Zhao Y, Lofi C, Hauff C (2018) Towards real-time webcam-based attention tracking in online learning. In: ACM annual meeting of interactive user interfaces (IUI)
25. Salazar J (2010) Staying connected: Online education engagement and retention using educational technology tools. *American Society for Clinical Laboratory Science* 23(3 Supplement):53–58
26. Sedrakyan G, Mannens E, Verbert K (2019) Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *J Comput Lang* 50:19–38
27. Stott P (2016) The perils of a lack of student engagement: Reflections of a “lonely, brave, and rather exposed” online instructor. *Br J Edu Technol* 47(1):51–64

28. Heckel T (2013) Canvas-Based Chart.js Version 0.1 Released Accessed from <https://www.infoq.com/news/2013/03/chartjs-v.0.1-released>
29. Wang MT, Degol J (2014) Staying engaged: Knowledge and research needs in student engagement. *Child Develop Perspect* 8(3):137–143
30. Ward MO, Grinstein G, Keim D (2010) Interactive data visualization: foundations, techniques, and applications. CRC Press
31. Whitehill J, Serpell Z, Lin YC, Foster A, Movellan JR (2014) The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Trans Affect Comput* 5(1):86–98
32. Xin C (2009, December). E-learning applications and challenges. In: 2009 Second international conference on future information technology and management engineering, IEEE, pp. 580–583



# Customers' Willingness to Use New OTT Media Platforms in India



Imon Chakraborty , Ravinder Kumar Verma ,  
Mohit Hiteshkumar Joshi, and P. Vigneswara Ilavarasan

**Abstract** Digital platforms have transformed businesses. Over-The-Top (OTT) platforms, for example, Netflix, Amazon Prime, Google enter the entertainment industries. As a result, existing media players imported this concept to develop their own OTT Media Platforms (OMP). With the supportive environment, OMP has grown many folds for instance, in India, there are almost 40 different OTT platforms. OMPs are competing for new markets, new customers, technologies, and sustaining existing businesses. Therefore, to sustain and grow in this competitive field, OMP should understand customers' preferences and factors to grow businesses. This study investigates the customers' Willingness to Use New OTT Media Platforms (WUNO). The study is based on the survey results of 1310 respondents from 11 Indian states. Willing to use new OMP (WUNO) is influenced by the content, subscription, age, language, time spend on OMP, and current usage on OMP. Individual choices and OMP features characterize the current usage and WUNO. Usage is preferred when OMP features are presented as a combination of the WUNO model and personalized as per the customers. The study findings are useful for the existing and new OTT platforms. It is contributing to the literature of OTT media platforms.

**Keywords** Digital platforms · Over-The-Top · OTT · Media platforms · Consumption · Willingness to Use · Entertainment content

## 1 Introduction

With the development of mobile technology and the expansion of the internet, users can watch a diverse range of digital media content at any time and space on smart devices. The outgrowth of internet-based digital media platforms has brought a

---

I. Chakraborty (✉)

UQ-IITD Academy of Research, Indian Institute of Technology, Delhi, India

e-mail: [imon.chakraborty@uqidar.iitd.ac.in](mailto:imon.chakraborty@uqidar.iitd.ac.in)

R. K. Verma · M. H. Joshi · P. V. Ilavarasan

Department of Management Studies, Indian Institute of Technology, Delhi, India

paradigm shift in the broadcasting industry. This emerging media platform is recognized as ‘Over The Top’ (OTT). OTT is delineated as the distribution of digital media content using the internet over a public network [1]. The communication service regulator in the US, Federal Communications Commission, defined OTT as “any entity that offers video content by means of the Internet or other Internet Protocol (IP)-based transmission path provided by a person or entity [2]”.

Compared to traditional broadcasting, the OTT Media Platforms (OMP) have lot of flexibility and variety, such as smartphone app, websites, and social networking services. This service is compatible with traditional Televisions (TV) to Personal Computers, including smartphones, and provided by any third party, including traditional broadband players. Therefore, the market is highly competitive with various multinational service providers such as Amazon Prime, Hotstar, Netflix [3]. OTT services have grown apace globally. In India, the OTT entertainment market is growing with a Compound Annual Growth Rate of 30% [4]. According to a report by BCG, this market will reach \$5 Billion by 2023 [5].

The OMP usage is depending on number of factors like subscription, content, supporting infrastructures. The OTT services offer three ways: one is a paid OTT service where users can only enjoy content if they have paid the monthly subscription fees; the second is entirely free, but it is an advertisement-based subscription. Viewers need to bear with some ads in between the shows they are watching. The third one is the Freemium model, in which some content are free, and premium content are available via paid subscriptions only [6]. Besides, OMP’s must inquire about consumers’ preferences and demands to establish a business strategy for flourishing maximally in the market. The existing literature provides very limited evidence to direct the emerging OMP’s in this regard.

Given that, this research is exploring the Willing to Use New OTT Media Platforms (WUNO). The results provide the factors that influence the customer’s decisions to use or not to use new OTT platforms. The study findings might assist the OTT platforms, practitioners, and academicians for best practices and decision-making towards better consumer acceptance.

The paper is divided into six sections. The first section briefly introduced the paper. The second section gives an overview of the literature on the OTT media platforms and theoretical lens. The third section talks about the research design and methodology. The fourth section depicts the findings, and the fifth section discusses those findings. The sixth section concludes the paper with implications and limitations.

## **2 Literature Review**

### ***2.1 Usage of OTT Media Platform***

The existing studies have focused on the impact of OTT platforms and their effectiveness compared to conventional TV platforms. Specifically, a few previous studies

analyzed the quality of services and consumer experiences, including the future of OTT media, and speculates the OTT service market [7].

The impact of OTT platforms on conventional media services from consumers' perspective have attracted researchers' attention. A study revealed that the OTT platform, YouTube has maximum popularity and satisfies the users in all dimensions. Besides YouTube, local OTT media services attract a maximum number of consumers because of the relevance of its content according to the viewer's needs [8].

A few studies examined the consumers' usage behavior and preference of emerging platform's services. Usage patterns highlight the consumers' choice about the content and their time spent on it. Quresh Moomchala [7] investigated the content preference along with the influence of languages on the OTT media contents. The study also analyzed the driving factors for internet users towards OTT digital streaming.

Based on the existing studies, the impact of age group comes just after the content availability and its type [9]. Many studies reported that the usage, content preference, subscription, and willingness to use the OMP significantly vary with the consumer age. The young adults have shown a prime interest in OTT platform adoption. However, subscription depends on the income bracket of the OTT platform user [10]. One study captured India's millennial consumers to understand the adoption factors of OTT media services [11]. Besides, an empirical analysis of the telecommunication services has reported a surge in OTT platforms due to the impact of the internet and mobile phone compatible media services [12].

Nevertheless, past literature empirically examined user behavior towards OTT services and their effectiveness over traditional media services. To the best of our knowledge, existing research does not focus on investigating the factors that direct the emergence and willingness to use new OTT platforms in developing countries.

## ***2.2 Theoretical Background***

The critical aspects of a new OTT platform should also be considered for their survival in a competitive market. This has not been investigated extensively. Moreover, with respect to the inconsistent results in terms of factors, the 'survival of the fittest,' a concept of Charles Darwin's evolutionary theory [13] is explained in this section, and it further assists in conceptualizing the study. According to the theory, the evolution mechanism can be conceptualized with Variation, Inheritance, Selection, and Time (VIST) [14]. In biology, Variation is a result of mutation of copying genetic material; Inheritance is passing the genetic code to offspring; Selection is survival in the environment, and Time denotes the survival of advantageous trait generation after generation.

In line with the emerging platform's survival, variation should be in terms change. The change in the content genre, number of services, area of operations, technological upgradation, and research in technology-based entertainment content development and promotion. The new changes in digital platforms-based media propagation built

**Table 1** Factors for current OMP usage

S. no	Factors	Literature
1	Free subscription	Li [17], Prince and Greenstein [18]
2	Paid subscription	Madnani [9]; Dasgupta and Grover [19]
3	Content	Li [9]; Madnani et al [17]
4	Age	Madnani [9]; Moochhala [7]
5	Time	Madnani [9]; Kim et al. [8]
6	Language	Moochhala [7]

over and along with the basic traditional TV or OTT platform's services. OMP inherited the traditional TV features and entertainment content so that consumers have multiple entertainment contents using the OMP. Selection is the ability to capture the maximum number of consumers in the competitive market. The platform needs to inculcate the magnitude for long-term survival.

Using the concept of evolutionary theory, 'survival of the fittest,' enables the authors to examine the factors to direct the new OTT platform for their survival in the competitive market. In existing information systems literature, the theory has been used for the evolutionary perspective of digital platforms usage, understanding of e-communication behavior, adaption of e-communication, and so on [15, 16]. However, the survival perspective is not explored yet for the OTT media platform. The present work aims to use that theoretical lens to examine the fittest aspects of the new OTT media platform.

### 3 Methodology

#### 3.1 Measurement Instrument

The variables of the questionnaire are taken from the literature on OTT Media entertainment and digital platforms. The select variables' sources are listed in Table 1. The variables are contextualized according to the study objectives and Indian context.

#### 3.2 Data Collection and Analysis

Data were collected using the survey tool between 15th June 2018 and 10th July 2019. A structured questionnaire was developed in English, which comprises both closed-ended and open-ended responses. The feedback from the pilot study has been considered for reforming the questionnaire.

The study respondents comprise of following age group customers above 15 years, using the first age group bracket as 15–21 years and last customers above 60 years. A total of 85,000 respondents were contacted online, out of which, survey was forwarded to 45,000 using their emails and 40,000 customers through text messages on their mobile number. The study receives 1310 complete responses after removing incomplete responses. The survey data covers 11 Indian states though majority data come from two states Gujarat and Maharashtra, where most of the responses are coming from major metro cities.

The data have been analyzed using the frequency distribution using the SPSS statistical tool. The questionnaire survey captured into a Microsoft excel sheet and imported into SPSS software for statistical analysis.

4 Findings

4.1 Willingness to Use New OMP

The current OMP subscribers are more likely for WUNO. Figure 1 shows that 79% of the users who are OMP subscribers are willing to use the new OTT apps, whereas 67% of people who are currently not subscribing to any of the OTT apps are willing to use the new OTT app. The subscription influences the users’ WUNO, this reflects the users’ preference towards the freely available features of the new OMP.

The paid OMP subscribed users showed higher WUNO compare to free subscribed users of OMP. Figure 2 shows WUNO interests of users, 82% of the users with paid OMP are interested and 71% of the users having a free OMP subscription.

The spatial content influences the users’ interest in WUNO. Those users who like the regional content and they prefer WUNO compare to those users who have less preference for regional content. Figure 3 shows that 86% of the people who are having preferences for the regional content are WUNO, whereas 59% of users with less preference for regional choose for WUNO.

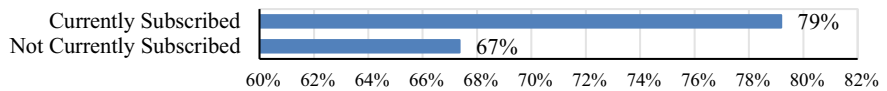


Fig. 1 Current OTT subscriber versus willingness to use the OTT app

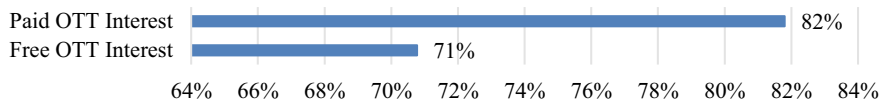
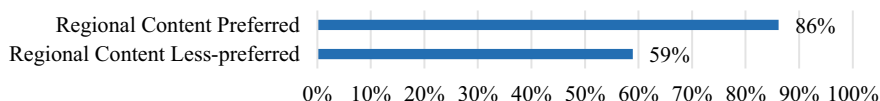


Fig. 2 Paid OTT interest versus willingness to use the OTT apps



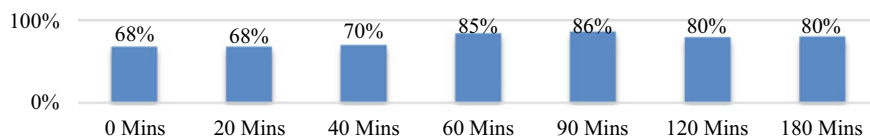
**Fig. 3** Regional content preference versus willingness to use the OTT apps

It is showing that the user's attachment with OMP is important in terms of the user's preferred content. The regional content engages the users with the digital platforms, this strengthens the OMP viewing behavior of users. The time spend on OMP is directly related to the WUNO as shown in Fig. 4. Those users who spend 90 min or higher have higher preference for the WUNO.

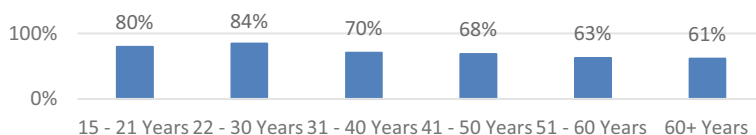
The users age influences the preference for WUNO as shown in Fig. 5. The users age is directly related to the preference for WUNO, lower the age of the users higher the preference for WUNO. The users in the age group of 15–21 years where 80% of the users are WUNO and 84% of users in the 22–30 years and subsequent age groups preference for WUNO percentage is decreasing.

To find the relationship between the language and users WUNO, analysis shown in Fig. 6. The preference of users for a particular language is reflected as a medium rather than the preference, as most of the users' preference for their mother tongue shows a higher WUNO for all languages with little difference.

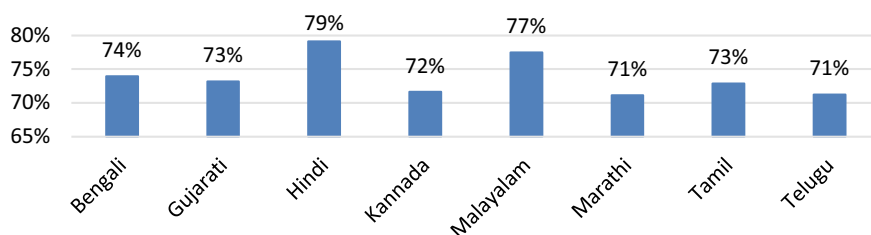
However, data show that the users prefer regional content and mother tongue as a medium of entertainment; however, mother tongue has less impact on WUNO. The limitation of the mother tongue is overcome through technology by making entertainment content available in the popular language of a nation.



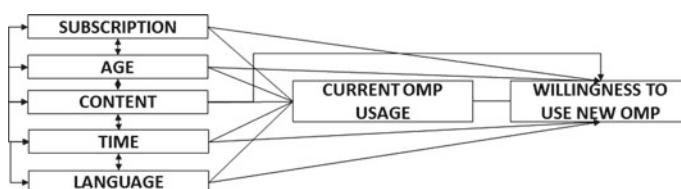
**Fig. 4** Daily OTT time spent versus willingness to use the OTT apps



**Fig. 5** Age group wise willingness to use the new OTT applications



**Fig. 6** Mother tongue wise willingness to use the OTT apps



**Fig. 7** Conceptual model of OTT media platform usage

## 4.2 Conceptual Model of OTT Media Platform Usage

The WUNO and OMP usage is influenced by the multiple variables, this shows that the users' preferences for the packaged digital platforms with varieties in OMP characterization. The conceptual model can be divided into three parts as shown in Fig. 7. The first part discusses five variables influencing the usage behavior, and the second part highlights the variables influencing the current OMP usage and WUNO. The third part is the interactions between the current usage and WUNO, including the interrelationship between variables.

The most preferred OMP is generally a combination of individual needs and packages offered by the OMP. The current subscription and usage preferences are impacting the WUNO.

## 5 Discussion

The usage of OMP and WUNO has shown a trend towards the bundle of entertainment package. OMP has moved away from the traditional media industry towards customer centric approach this approach is guided by competition and remain relevant. These aspects decide the survival and growth of the OMP.

The variance in the choices of customers led to multiple content genres on OMP. Consumers' usage is related to subscription preference, content, user age, daily time spent on OMP, and language. The outcome shows that current subscribers are more prone to WUNO than new subscribers. An empirical study by Li [17] showed that

consumers were mostly engaged with YouTube because of its no subscription charge policy [17].

Customers inherit the past practices of subscription, non-subscription, and content preferences of OMP. It reveals the preference of customers for free-content platforms; however, complementary and freemium based services have a strong potential to attract the consumer [10]. Teenagers are among all age groups for any subscription are more inclined toward the free OMP services, but the young adults are more predisposed to paid services among all [20]. Considering the factors for survival of the fittest in a competitive market, new OMP should target different age groups, provide specific content- regional content, live content, carrying over of past popular practices such as television serials, and language support. Taiwan's based study on the OTT platform reported that serials, movies, and sports are most desirable among users [21].

New OMP needs to stress on the subscription model, which must also include a freemium-based approach. If an emerging platform considers these key attributes, then the evaluation mechanism, VIST, will be strengthened for value creation in the competitive market. A vast content genre focusing on regional content and language can make the variation by reaching any age group even beyond the preferences of a particular age group. Only a single study revealed the impact of language on OTT services [7].

Inheritance is captured by considering traditional television content and language. Selection is achieved when the new OMP customizes the subscription plan according to the user's usage behavior. This links with the time concept, long-term survival, including value capture in the market by incorporating free and freemium strategy, considering the consumer's time spent on OMP. Viewing through these theoretical lenses, the study depicts the conceptual framework with the components such as free subscription, paid subscription, age, content, time, and language, which impacts the current OMP usage and WUNO adoption.

## 6 Conclusion

The present research talks about the consumer willingness to use new OMP services in India. It depicts those users significantly favor live broadcasting services, including regional content. In terms of subscription, free OMP has the highest popularity, albeit paid users are more interested in new OMP to trade off their subscribed services.

We have shown the use of evolutionary theory's VIST in the analysis of the WUNO. This study shows the shift of the entertainment industry to digital platforms and reformatting of the media content as per the customers preferences. It bridges the existing literature gap by considering the theoretical view and strengthening it using an adequate sample size. Additionally, by introducing a conceptual model, the study directs the OMP business practices and opens the opportunity for further research.

The present study is not without limitations. The study only considered the variables extensively found in the literature for OTT media usage. Therefore, a further



deeper investigation of consumers from various income groups, varying demographics, and lifestyles can produce more insights about the OMPC in the developing country context. Due to space constraints, this paper presents a part of the entire study.

The present paper contributes on theoretical and practical fronts. Theoretically, this work shows the usage of an evolutionary framework in digital platforms for survival context, but more in-depth theoretical grounding is desired. Nevertheless, for practitioners, this study provides insights into the establishment and continuance of OMP. In the initial phase, OMP should follow the freemium model, with region and age specific content, in long terms OTT platforms evolve to have bundle of media content to serve multiple users preferences. The study intends to further investigate and nullify the falsifiable hypothesis by more theoretical underpinning with empirical data. Therefore, the work encourages more empirical studies that explore the evolutionary theory and survival of the fittest concept to test the survival aspects of OTT media platforms.

## References

1. Gonçalves V, Evens T, Alves AP, Ballon P (2014) A service of zbw power and control strategies in online video services
2. Report FCC-NA, MB, FD (2004) Undefined: annual assessment of the status of competition in the market for the delivery of video programming
3. Shin J, Park Y, Lee D (2016) Strategic management of over-the-top services: focusing on Korean consumer adoption behavior. *Technol Forecast Soc Change* 112:329–337. <https://doi.org/10.1016/j.techfore.2016.08.004>
4. OTT Platforms Report. <https://www.mica.ac.in/indian-ott-platforms/indian-ott-platforms>
5. Naik NV (2021) An introduction to over-the-top entertainment. India—a perfect playground for this digital industry. Presented at the May 14
6. Doorsanchar M, Lal BJ, Marg N (2015) Regulatory framework for over-the-top (OTT) services
7. Mochhala Q (2018) The future of online OTT entertainment services in India
8. Kim J, Kim S, Nam C (2016) Competitive dynamics in the Korean video platform market: traditional pay TV platforms vs OTT platforms. *Telemat Info* 33:711–721. <https://doi.org/10.1016/j.tele.2015.06.014>
9. Madnani D, Fernandes S, Madnani N (2020) Analysing the impact of COVID-19 on over-the-top media platforms in India. *Int J Pervasive Comput Commun* 16:457–475. <https://doi.org/10.1108/IJPC-07-2020-0083>
10. Fitzgerald S (2019) Over-the-top video services in India: media imperialism after globalization
11. Grand T, Writer F, County F (2019) Over-the-Top Video Services Among 10:61–71
12. Woo J, Choi J, Shin J, Social JL-T, forecasting U (2014) The effect of new media on consumer media usage: an empirical study in South Korea. Elsevier
13. Gould S (1982) Darwinism and the expansion of evolutionary theory. *science.sciencemag.org*.
14. Spiegel A, Evans M, Gram W, Diamond J (2006) Museum visitors' understanding of evolution. *Museums Soc Issues* 1:69–86. <https://doi.org/10.1179/msi.2006.1.1.69>
15. Kock N (2002) Evolution and media naturalness: a look at e-communication through a darwinian theoretical lens. *ICIS 2002 Proceedings*
16. Ng YL (2019) Toward an evolutionary perspective on social media use for cooperation. *Evol Behav Sci*. <https://doi.org/10.1037/ebs0000172>

17. Li S (2017) Television media old and new: a niche analysis of OTT, IPTV, and digital cable in Taiwan. Elsevier
18. Prince J, Greenstein S (2017) Measuring consumer preferences for video content provision via cord-cutting behavior. *J Econ Manag Strateg* 26:293–317. <https://doi.org/10.1111/jems.12181>
19. Dasgupta S, Grover P (2019) Impact of digital strategies on consumer decision journey. *Acad Mark Stud J* 23(1):1–14
20. Policy YC-T (2019) Undefined: competitions between OTT TV platforms and traditional television in Taiwan: a Niche analysis. Elsevier
21. Katherine Chen YN (2019) Competitions between OTT TV platforms and traditional television in Taiwan: a Niche analysis. *Telecomm Policy* 43:101793. <https://doi.org/10.1016/j.telpol.2018.10.006>

# Application of Java Relationship Graphs to Academics for Detection of Plagiarism in Java Projects



Divya Tyagi , Ritu Arora , and Yashvardhan Sharma 

**Abstract** In today's online learning environment, plagiarism detection tools are increasingly used by teachers and instructors to restrain students from plagiarism. Moreover, need for plagiarism detection tools to detect plagiarism in programming assignments has also increased manifolds. In this paper, we present the application of Neo4j Graph Databases to detect similarity between Java program submissions by students, in academics. This is done by converting a Java program into a specialized dependency graph and then implementing various comparison techniques on this graph. The two graph comparison techniques proposed and implemented in this paper are based on structural comparison of graphs by node-type count comparison and elemental comparison of method nodes in graphs by body-element-count comparison. The results of these two techniques are combined with the call graph-based technique, proposed in an earlier work, to calculate overall similarity index between program codes. This study captures a large category of changes that may be introduced to the code for plagiarism.

**Keywords** Plagiarism detection · Computer science · Object oriented programming · Java programs · JavaRelationshipGraphs (JRG) · Dependency graphs · Call graphs · Neo4j graph databases

## 1 Introduction

With the advancement of technology and the Internet, it has become easier for students to copy assignments from their peers. Plagiarism has various negative repercussions. Firstly, a student can take the credit for the work not originally produced by him/her. Moreover, by plagiarizing someone else's work, the student gets no knowledge of the subject matter. And this, in the long run, undermines the validity of

---

D. Tyagi (✉) · R. Arora · Y. Sharma  
Birla Institute of Technology and Science, Pilani, India  
e-mail: [f20170727@pilani.bits-pilani.ac.in](mailto:f20170727@pilani.bits-pilani.ac.in)

Y. Sharma  
e-mail: [yash@pilani.bits-pilani.ac.in](mailto:yash@pilani.bits-pilani.ac.in)

academic degrees. According to the survey conducted by Donald McCabe, Rutgers University (2005) [1], 70% of total students surveyed (over 70, 000 US graduate and undergraduate students) were found to introduce plagiarism. Not only in the universities, but the plagiarism cases are common in companies also. These cases are graver and concerning, as they may even cause court disputes. However, our focus is primarily on detection of plagiarism in academics. Plagiarism detection has become even more important in today's online learning environment. Due to the ongoing pandemic situation, more and more schools and universities have been forced to shift to the use of various online platforms for teaching course content, and for conducting and evaluating laboratory assignments. As a result of this plagiarism cases have increased tremendously. So, detecting plagiarism is important and tedious job on the part of instructors.

## 1.1 Objective

The main aim of this research is to detect similarity between Java source codes. The objectives of this work are—converting Java programs into specialized dependency graphs (based on previous work), detecting plagiarism using structural and elemental comparison of graphs and devise a powerful similarity index to detect plagiarism.

For this study, we use JavaRelationshipGraphs (JRG) [2, 3] stored using Neo4j Graph Databases. JRG is a specialized dependency graph and Java program code can be easily converted into its equivalent JRG using the JRG framework [3]. Moreover, this study is complementary to an already existing graph-based technique of detecting plagiarism [4]. In the previous study [4], call graphs were obtained from JRG and further used to calculate the similarity index between program codes. The motivation is to further enhance the results of the previously conducted research. In this paper, we present the novel approach by combining the results of the call graph-based technique with the structural and elemental comparison techniques, proposed in this work. Our goal is not to find true definite cases of plagiarism, but we are more focused on cases that are suspicious enough so that a teacher should assess them. The proposed technique of structural comparison of graphs by node-type count comparison, counts and matches the number of nodes of a certain type and uses this count to compare the structure of graphs obtained from their corresponding Java programs. Moreover, the elemental comparison of method nodes in graphs by body-element-count comparison technique, counts and matches various basic elements existing in a Java program body like number of lines, blocks and keywords (if, while, for, new, etc.).

The remaining paper is structured as follows: Sect. 2 discusses the relevant works in the field of plagiarism detection while Sect. 3 presents the proposed techniques. In Sect. 4, the proposed method is demonstrated using an example and results are discussed. And finally, Sect. 5 concludes the paper with a brief note on future work.

## 2 Related Works

With the advancement of technology, different approaches have been used by researchers, including text and token-based comparison [5], graph-based [4, 6–9] and tree-based techniques [7]. In [10], a token sequence is generated using Java class files and then similarity is evaluated using adaptive local alignment. In CAPlag [11] attribute-based method is combined with structural based method and hidden structural similarities are analyzed in a program along with evaluating important attribute-based metrics.

In graph-based techniques, dependency graphs are very effective in representing the relationship between different program elements in the source code. Static call graphs technique [4] that is based on application of JRG [3] represent the method calls between different methods in static source code. In [6], the author extracted the call graph from Abstract Syntax Tree of program source code and then used n-grams technique. Hyun [7] used a combination of syntactic and dynamic information to compare programs. For syntactic information, parse trees are compared using specialized tree kernels, while for retrieving dynamic information, call graphs are studied for isomorphism. In Program Interaction Dependency Graph [8], the behavior of programs in terms of data used and how it interacts with the system is represented. BPlag [9] analyses the two programs based on their execution behavior and represent a program in a graph-based format.

Some of the most popular existing detection tools are: -

- MOSS [11]—It is a freely available tool. Information of the algorithm used, and test results are not provided for MOSS. It detects plagiarism in programming languages—C, ML, Pascal, C++ , Java, Scheme, Lisp, and Ada.
- JPLAG—It works for the programs written in C, Java, Scheme, C++ , free text languages. Test results are not published for JPlag. However, it makes use of the Greedy String Tiling (GST) [12] technique to compare pairs of strings for evaluating similarity and it also provides a robust graphical interface.

## 3 Proposed Method

In this work, we extend the usage of JavaRelationshipGraph (JRG) framework [3] to academics for detection of plagiarism in Java projects. The framework converts a Java project into a dependency graph, comprising of the seven program elements that exist in any Java program, namely—project, package, class, interface, method, class-attribute, and method-attribute. Each program element is connected to its owner element using ownership relationship. Additionally, a Java project might consist of five different types of dependency relationships—extends, implements, imports, uses and calls. Moreover, a JRG is created using Neo4j Graph Databases. Graph databases are based on property-graph model and consist of entities, called nodes, which are connected to other nodes through directed relationships, called edges. Both

nodes and edges can have any number of properties (key-value pairs) associated with them. In a JRG, nodes represent the program elements and edges represent the ownership and dependency relationships. Further, we convert Java projects submitted by students, into corresponding JRG and apply various algorithms to compare these graphs, calculate similarity scores and hence detect plagiarism. One such technique was proposed in our previous work [4], in which we obtained the call graph from the JRG and then compared the call graphs of various submissions and obtained a measure of similarity score. In this work, we add two more techniques for comparing graphs obtained from Java projects and calculate corresponding similarity scores. Finally, a combined similarity score is calculated, which combines the similarity score obtained from all the three techniques. This helps us achieve a more effective similarity score, better accuracy, and less false positives. The sub-sections below discuss the two new proposed techniques.

### 3.1 *Structural Comparison of Graphs by Node and Edge Count Comparison Technique*

In the first technique, we carry out structural comparison of JRGs created from Java programs by counting the nodes and edges of the same type. In this study, we count only five type of nodes—class, method, interface, class-attribute and method-attribute nodes. Project and Package nodes are not counted because they mostly remain the same. Students usually do not make changes to these higher-level nodes while introducing plagiarism. Additionally, while counting edges, ownership edges are not counted because these edges connect nodes together and hence are indirectly counted while counting the nodes. Moreover, only three types of dependency edges are counted—calls, uses, and imports. Dependency edges of types—extends and implements are not counted for the same reason as discussed above for project and package nodes.

After calculating these nodes/edges for the two JRGs, a similarity % between them is calculated using the exponential weighted difference formula which is given by: -

$$\text{Structural Similarity \%} = (e^{-\sum_{i=1}^8 (\text{weight})_i |\text{difference}|_i}) * 100 \quad (1)$$

where *difference* is the subtraction of corresponding count values of nodes and edges between the two graphs. If the difference in nodes/edges counts (for corresponding node/edge type) is zero, this implies there are high chances of code being plagiarized because the structure is similar. Absolute value of difference is taken to avoid the effect of nullification. And *i* moves from 1 to 8 since there are a total of eight elements (five node and three edge types). The points that were kept in mind while devising the formula are—the function used should be a decreasing function because

**Table 1** Weights corresponding to different nodes and edges

S. no	Node/Edge type	Weight
1.	Interface nodes	0.05
2.	Class nodes	0.04
3.	Method nodes	0.02
4.	Class-attribute nodes	0.01
5.	Method-attribute nodes	0.01
6.	Calls edges	0.02
7.	Uses edges	0.03
8.	Import edges	0.01

if summation of weighted differences is small, then similarity % should be high. And the mathematical function range should be (0,1). Therefore, an exponential function is considered.

For deciding the weights, the criteria considered is if the element is easier to modify or more prone to plagiarism modifications, then its corresponding weight is kept low, because it's significance in the similarity percentage should be less. We can understand this as it is less likely that students will introduce dummy classes/interfaces to make plagiarism changes, so class and interface nodes are given more importance. While a student tries to add dummy attributes/methods or break methods into smaller, therefore method, attribute and method-attribute nodes are given less weight. Weights for dependency edges are also decided in a similar manner. The weights are adjusted manually by analyzing similarity % after introducing small plagiarism changes in programs to decide final weights. The weights thus obtained are given in Table 1.

The plagiarism changes detected by the proposed structural comparison technique are:

- Plagiarism changes introduced to comments, white-spaces, and code formatting—since these elements are not considered in the counting process, so any changes made to these do not cause changes to the similarity %.
- Plagiarism changes introduced at Class/Interface level
  - Renaming or changing the signature of class(es)/interface(s)
  - Rearranging attributes and methods within a class/interface
  - Addition or deletion of classes/interfaces
- Plagiarism changes introduced at Method level
  - Renaming or changing the signature of methods
  - Reordering code blocks within a method
  - Addition or deletion of method(s)
  - Breaking a method into smaller methods
  - Replacing program statements with equivalent semantics in a method
  - Modifying expressions and changing decision logic within a method

- Plagiarism changes introduced at Attribute level for class and methods
  - Renaming or changing the data types of attribute(s)
  - Addition or deletion of attribute(s)

If the student has just changed the names or signatures of classes, interfaces, methods or/and attributes, then the weighted difference of nodes/edges obtained will be zero. Similarly, rearrangement of methods and attributes within a class and reordering code blocks within a method results in a weighted difference which is also zero. Changing method statements like replacing if statements with semantically equivalent switch cases or modifying expressions does not change the number of methods in the code. Therefore, any such changes deliberately introduced in the code will also be detected by our technique. The reasoning for Interface level changes will be the same as that of class level changes. If the student tries to break existing methods into smaller ones, then most likely he/she will break it into two. This will increase both method nodes and calling edges by one. But as weights corresponding to both are low, therefore similarity % does not get affected drastically. And if dummy methods are created, then also this will increase method count but method calls remain the same, hence it does not also change similarity % significantly. Similar is the case for introduction of dummy attributes and classes. Deletion of class, interface, method, and attribute has the same effect as that of addition, because the absolute difference of count will remain the same.

### ***3.2 Elemental Comparison of Method Nodes in Graphs by Body-Element-Count Comparison Technique***

In this technique, the JRGs obtained from the program submissions are compared for elemental similarities of the basic elements that exist within the method-body for corresponding method nodes. Importantly, while creation of a JRG, the method-body is stored as a property of the method node, in the form of key-value pair and be easily retrieved. Moreover, different elements in the method-body are counted and then the overall similarity % is calculated. The basic elements used are:

- Semicolons—to count the total number of lines in the method-body
- Curly brackets—to count the total number of blocks in the method-body
- Full stop—to count the total number of function calls
- Keywords—to count various elements in the method-body; keywords considered here are:—break, continue, try, catch, for, while, if, else, switch, case and new.

Moreover, as generally students rename the methods to introduce plagiarism, so to find out which method node in one graph is to be compared with which method node in another graph, three things are matched in the two methods - parameter list, return type and modifier. And the exponential weighted difference formula used to calculate similarity percentage between two method nodes is given by:



**Table 2** Weights corresponding to various elements of the method-body

S No	Description	Weight
1.	<b>break</b> keyword	0.03
2.	<b>continue</b> keyword	0.03
3.	<b>try</b> keyword	0.03
4.	<b>catch</b> keyword	0.03
5.	<b>for</b> keyword	0.02
6.	<b>while</b> keyword	0.02
7.	<b>if</b> keyword	0.02
8.	<b>else</b> keyword	0.02
9.	<b>switch</b> keyword	0.02
10.	<b>case</b> keyword	0.02
11.	<b>new</b> keyword	0.02
12.	Blocks	0.02
13.	Function calls	0.01
14.	Lines	0.01

$$\text{Elemental Similarity \%} = (e^{-\sum_{i=1}^{14} (\text{weight})_i |\text{difference}|_i}) * 100 \quad (2)$$

Here, *weight*, *difference* and *i* have the same meaning as in Eq. (1). We are considering total 14 elements so *i* moves from 1 to 14. For deciding the weights, the same criteria are adopted as in Eq. (1). For instance, since it is less likely that a student changes break, continue, try and catch expressions, therefore they have more weights. While it is quite easy to alter the number of lines by simply introducing dummy print statements; hence given less weight. Table 2 lists the weights assigned to various elements, in increasing order of difficulty in modifying the element—which also leads to decreasing order of weights. The easier it is to change an element, the lower the weight assigned to it.

The steps involved are detailed in Algorithm 1. If a student copies the entire code or changes the comments and indentations or renames the methods or rearranges the attributes, then the similarity % calculated will be 100%. Moreover, if a student changes statements like replacing for-loop with semantically equivalent while-loop, this will not change the similarity % significantly as weights associated with these keywords are very low. Moreover, if a student tries to break a method into more methods, it will increase the number of blocks and parenthesis, which again have less weightage, while other element count will remain same; therefore, similarity % will not change drastically. Additionally, if there is introduction of dummy method(s), then these method node(s) in one graph will not correspond to any method node(s) in another graph, so will be dropped without affecting the similarity %.

Algorithm 1: To calculate Elemental level-based Similarity Percentage

Input	JRG1 and JRG2—Java Relationship graphs of two Java Projects
Output	Similarity Percentage
Step 1	Run the CQL query on the JRG to obtain method name, return type, parameter list, modifier, and method-body of all the method nodes for both JRGs
Step 2	Create an array list of methods and store the details obtained from Step 1 as the attributes of Method object
Step 3	Calculate the counts of above-mentioned elements with respect to each method and store in an array list
Step 4	To find pairs of most similar methods in two graphs, compare parameter list, return type and modifier of every combination of method node pairs
Step 5	Calculate the similarity between pairs of most similar method nodes obtained in Step 4, using the exponential formula designed for this technique
Step 6	To obtain overall similarity score, find the average of similarity score of different pairs of method nodes calculated in Step 5

### 3.3 Combining Results with Call Graph Based Technique

In the call graph technique [4], the author converts each program into its JRG and then extracts its call graph from it. Then in the next step, compares call graphs of two program for isomorphism and calculates similarity index. The formula used to calculate the similarity percentage is given by:

$$\text{Call Graph Similarity \%} = \left( \frac{\text{Number of common elements}}{\text{Total number of elements in the arraylist}} \right) * 100 \tag{3}$$

This technique has higher chances of false positives, if considered standalone. Changes like addition of new method(s) by breaking a method into small ones, or introduction of additional dummy or useful attributes, are few types of changes that cannot be captured by the call graph technique. Therefore, in this work, we improvise on our results by extracting more structural and elemental features from the JRG of a program. After calculating additional similarity indices, we combine our results with the results of call graph technique. The average of all three is computed for overall similarity index.

For a large submission set of Java projects, each project is compared with all the other projects and an overall similarity % is calculated. Finally, the similarity % for a single project is calculated as the maximum of all the similarity scores obtained by comparison with all the projects. The threshold similarity percentage beyond which we will consider that code is plagiarized is 90%. This number is calculated experimentally.

The source code for the proposed technique is open sourced at: [https://github.com/DivyaTyagi1111/Plagiarism\\_Detection/blob/main/CommunityNeo4jMaven/src/org/jrg/parser/Structural\\_Elemental\\_Similarity.java](https://github.com/DivyaTyagi1111/Plagiarism_Detection/blob/main/CommunityNeo4jMaven/src/org/jrg/parser/Structural_Elemental_Similarity.java).

### 4 Experiment and Result

Let us consider an example Java program (Fig. 1). File EvenAndOdd.java is the original file which has three classes. Class *Even* contains a method *is\_even()*. Class *Odd* has one class attribute *value* and a method *is\_odd()*. Class *EvenAndOdd* provides the usage of the entire program. It contains the *main()* method which takes an input number and tells if it is even or odd. TestNumber.java is created by introducing changes: Some comments are added. All the classes and methods *is\_even()* and *is\_odd()* are renamed. A dummy class attribute named *dummy* is added in class *CheckEven*. The *if* expression is replaced with *else* expression in the *main()* method.

Figure 2 shows the results of running the proposed technique on the above code. JRG-1 and JRG-2 corresponds to EvenAndOdd.java, the original file and TestNumber.java, the plagiarized file, respectively. It depicts the similarity index by call graph, structural and elemental comparison techniques, and overall similarity % for the two JRGs. It also compares and displays the counts of different types of nodes and edges in the JRGs. The analogous method nodes in the two graphs are also presented. As the call-graph based similarity is 100%, so one may misinterpret that this is the best technique, but the fact is it does not consider many aspects of the

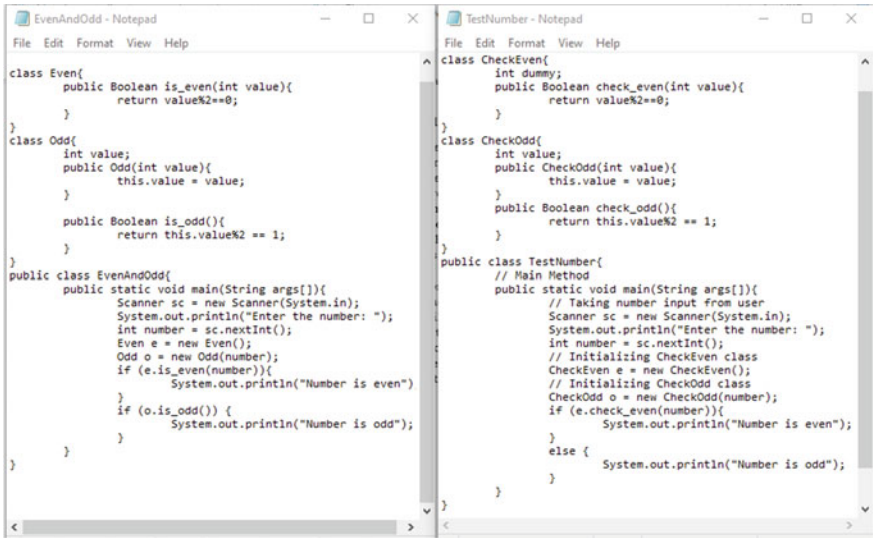


Fig. 1 Original and plagiarized code

```

Console
<terminated> ExtendedJRGCreator (2) [Java Application] C:\Users\DIVYA TYAGI\p2\pool\plugins\org.eclipse.justj.openj
***** Using Call Graph Technique *****
Similarity percentage at call graph = 100%

***** Using Structural Comparison Technique *****
          JRG-1  JRG-2
Class nodes =          3   3
Interface nodes =       0   0
Method nodes =          4   4
Class attribute nodes =  1   2
Method attribute nodes =  4   4
Calling edges =         2   2
Uses edges =           2   2
Import edges =          0   0
Similarity percentage at Structural level= 99.00%

***** Using Elemental Comparison Technique *****
is_even(int) <-> check_even(int)
Odd(int) <-> CheckOdd(int)
is_odd() <-> check_odd()
main(String) <-> main(String)
Similarity percentage at Elemental level= 98.78%

***** Combined Similarity *****
Overall Similarity percentage = 99.26%

```

**Fig. 2** Results of proposed technique

code as discussed earlier. The overall similarity % is 99.26% and as our threshold is 90%, so we accept the two codes to be plagiarized. Individually also, the similarity % obtained using the structural comparison technique and elemental comparison technique are above 90%. Even after introducing one attribute, replacing if with else expression, renaming classes and methods and adding comments, we get a high similarity %. Hence, the proposed technique is effective.

## 5 Conclusion and Future Work

This paper proposes a novel structural and elemental comparison-based approach for detecting plagiarism in Java programs. The technique makes use of dependency graphs generated from the Java program to make the overall plagiarism detection more robust. This technique uses the Neo4j graph database at its core. The proposed

technique and the associated algorithm calculate the similarity index between two programs. Using this technique, a large category of changes can be captured that may be introduced to the code for plagiarism. However, there exists a limitation of the proposed elemental comparison technique used to compare method-body for method nodes. It will give false results if there are overridden methods in the classes, since overridden methods have same signature. This can be overcome by prefixing the method names with the all the parent class names at the time of creation of JRG.

There are various directions for future work in this area. Firstly, the technique needs to be evaluated against diverse changes and real plagiarized program datasets. Secondly, as the proposed method deals only at the structural and elemental comparison level, it can also be combined with semantic and behavioral tools to enhance the results. Thirdly, different graphs like inheritance graph, control-flow graph, etc. could be extracted from JRG and used further to detect plagiarism. All these future works combined can help to investigate further if two programs are plagiarized or not.

**Acknowledgements** The authors would like to convey their sincere thanks to the Department of Science and Technology (ICPS Division), New Delhi, India, for providing financial assistance under the Data Science (DS) Research of Interdisciplinary Cyber Physical Systems (ICPS) Programme [DST/ICPS/CLUSTER/Data Science/2018/Proposal-16:(T-856)] at the department of CSIS, BITS, Pilani, India. Authors are also thankful to the authorities of BITS, Pilani, to provide basic infrastructure facilities during the preparation of the paper.

## References

1. McCabe DL (2005) Cheating among college and university students: a North American perspective. *Int J Educ Integr* 1(1)
2. Arora R, Goel S, Mittal RK (2016, August) Using dependency graphs to support collaboration over GitHub: the Neo4j graph database approach. In: 2016 ninth international conference on contemporary computing (IC3), IEEE, pp 1–7
3. Arora R, Goel S (2019) Java relationship graphs (JRG) transforming java projects into graphs using neo4j graph databases. In: Proceedings of the 2nd international conference on software engineering and information management, pp 80–84
4. Arora R, Maurya AM, Sharma Y (2021) Application of java relationship graphs (JRG) to plagiarism detection in Java projects: a Neo4j graph database approach. In: Proceedings of the 4th international conference on software engineering and information management (ICSIM), Yokohama, Japan, ACM. (Accepted)
5. Cavnar WB, Trenkle JM (1994, April) N-gram-based text categorization. In: Proceedings of SDAIR-94, 3rd annual symposium on document analysis and information retrieval, vol 161175
6. Steier M (2017) Discovering plagiarism in entry level programming assignments. <https://www.diva-portal.org/smash/get/diva2:1133441/FULLTEXT01.pdf>
7. Song HJ, Park SB, Park SY (2015) Computation of program source code similarity by composition of parse tree and call graph. *Mathem Probl Eng*. <https://doi.org/10.1155/2015/429807>
8. Cheers H, Lin Y (2020, January) A novel graph-based program representation for java code plagiarism detection. In: Proceedings of the 3rd international conference on software

- engineering and information management, pp 115–122. <https://doi.org/10.1145/3378936.3378960>
9. Cheers H, Lin Y, Smith SP (2021) Academic source code plagiarism detection by measuring program behavioral similarity. *IEEE Access* 9:50391–50412. <https://doi.org/10.1109/ACCESS.2021.3069367>
  10. Ji JH, Woo G, Cho HG (2008, November) A plagiarism detection technique for Java program using bytecode analysis. In: 2008 third international conference on convergence and hybrid information technology, vol 1. IEEE, pp 1092–1098. <https://doi.org/10.1109/ICCIT.2008.267>
  11. Menai MEB, Al-Hassoun NS (2010, August) Similarity detection in Java programming assignments. In: 2010 5th International conference on computer science and education. IEEE, pp 356–361. <https://doi.org/10.1109/ICCSE.2010.5593613>
  12. Wise MJ (1993) String similarity via greedy string tiling and running Karp-Rabin matching. *Online Preprint Dec* 119(1):1–17

# Predicting Appropriate Speed for Driving Based on External Factors



Adwait Bhope , Aashay Zangpure , Vinod Kamat , Rohan Dhamale ,  
Prashant Bhandari , Pruthi Joshi , Manisha Marathe ,  
and Urmila Kalshetti

**Abstract** The statistics of road accidents due to over-speeding are alarming. Therefore, there is a need for a robust system that will recommend a real-time safe speed limit to a vehicle driver. This paper proposes a feasible solution in the form of a hardware module that works on the principle of edge computing. The module estimates an ideal speed limit considering factors such as nearby vehicles, pedestrians and traffic signs as parameters. The speed is estimated on the basis of the distance between the vehicle and other vehicles/pedestrians and also on the basis of speed limit traffic signs. The estimated speed limit will then be notified to the vehicle driver. This will help in reducing the fatalities caused by over-speeding.

**Keywords** Accidents · Vehicles · Speed limit · Image recognition · Object detection · Edge computing

## 1 Introduction

Road accidents are increasing day by day, the prime cause being the over-speeding of vehicles. Excessive speed is responsible for one-third of all fatal accidents. As per the statistics shared by the Ministry of Road Transport and Highways—India, the total number of road accidents that took place in the year 2019 is around 4,50,000. Out of these, over-speeding accounts for the maximum share of road accidents (63–74%) and road accident deaths (62%).<sup>1</sup> Overspeeding by vehicles is a major issue for road safety and needs to be properly addressed to minimize accidents. This paper presents a potential solution that takes advantage of the advancements in technology.

---

A. Bhope (✉) · A. Zangpure · V. Kamat · R. Dhamale · P. Bhandari · P. Joshi · M. Marathe · U. Kalshetti

Department of Computer Engineering, Pune Vidyarthi Griha's College of Engineering and Technology Affiliated to Savitribai Phule Pune University, Pune, India  
e-mail: [mvm\\_comp@pvgcoet.ac.in](mailto:mvm_comp@pvgcoet.ac.in)

U. Kalshetti  
e-mail: [umk\\_it@pvgcoet.ac.in](mailto:umk_it@pvgcoet.ac.in)

<sup>1</sup> <https://morth.nic.in/road-accident-in-india>.

External factors, viz. traffic and pedestrian density, nearby places like schools and hospitals and traffic signs are analysed to calculate a safe driving speed that will be recommended to a driver. Adhering to this speed limit will prevent over-speeding, thus reducing accidents. The detailed approach is described in this paper and is organized as follows: Sect. 2 discusses related work, Sect. 3 gives an overview of the complete solution that this paper proposes, Sect. 4 describes the implementation details and the breakdown into smaller independent tasks, and lastly—the formula for calculating the appropriate speed. The simulation results are shown in Sect. 5, and Sect. 6 draws a detailed conclusion.

## 2 Related Work

Driver assistance systems can prove to be helpful in reducing accidents due to over-speeding and making the road traffic flow seamlessly. Various systems have been proposed and implemented to tackle the case of over-speeding and unnecessary traffic jams. In Frank et al. [3], the authors have put forth their idea of using a Raspberry Pi and a smartphone for recommending drivers a safe and reliable speed to avoid slowdowns and unnecessary traffic jams. It uses Raspberry Pi to obtain the speed of the vehicle and sends it to a smartphone which is then broadcast to all the drivers nearby using an access point. This method requires the installation of access points at regular intervals which is not cost-effective. The use of a smartphone makes the system vulnerable to power outage. The connection between Raspberry Pi and the smartphone is based on a USB interface which is clearly the driver's choice, which means that they might not use the solution at all times.

Sundar et al. [14] propose a system in which the road signals will be activated dynamically based on the vehicle density at that junction. They have considered that every vehicle has an RFID tag or a Zigbee transmitter for emergency vehicles and every junction has a reader. Based on the number of tags detected, the signals will be controlled dynamically to reduce the stopping time. Further, they also suggest that the use of such tags will help in tracking stolen vehicles. However, installation of readers and tags at every junction is not feasible. Also, there is a tampering and destruction risk associated with the tags. In Pathan et al. [11], the authors have implemented a solution using Raspberry Pi for the detection of speed, GPS coordinates and the driving lane. The authors have proposed the use of a client-server communication system in which the vehicle will report its speed, GPS coordinates, and the driving lane to a server. This data is then used to calculate the speed for the nearest vehicle driving in the same lane and sent to that particular vehicle. This system only considers the speed of the vehicle but does not consider the limit set by local authorities. Hosseini et al. [5] present a modelling approach to calculate a speed limit that minimized the cost incurred to the society in terms of road accidents. However, this is a one-time calculation and does not take into account real-time external factors such as traffic density.



### 3 Proposed Solution

The system design consists of four key components—a camera, an intelligent hardware chip such as Raspberry Pi or Jetson Nano, an integrated GSM-GPRS-GPS module and an inertial measurement unit (IMU) sensor. A camera is mounted in front of the vehicle and sends a live video feed to the Raspberry Pi. Regular cameras, as well as infrared cameras, are supported for better nighttime results. Várhelyi [15] proposes a solution for calculating an appropriate speed in low-visibility situations, but does not make use of an infrared dataset. The deep learning pipeline that this paper proposes runs on the Raspberry Pi and processes video in real time. It identifies vehicles, pedestrians and traffic signs. The GPS sensor helps in finding the type of road (street, highway, etc.), nearby places of interest, viz. schools and hospitals. These factors are fused to produce a single value which is then recommended to the driver as a safe speed limit. It is shown on a colour-coded LCD and sudden alerting warnings are communicated using an audio buzzer. The system starts automatically when the vehicle starts and requires zero inputs from the driver so that they can focus on driving.

The IMU is used to detect an accident when sudden jerks are sensed. These are reported to a server for analysis on accident-prone areas. The impact of this solution grows many-fold when this system is used by more and more vehicles. It enables the collection of data for accident-prone areas, which is another powerful factor for calculating a recommended speed.

## 4 Model

### 4.1 Data Gathering and Preprocessing

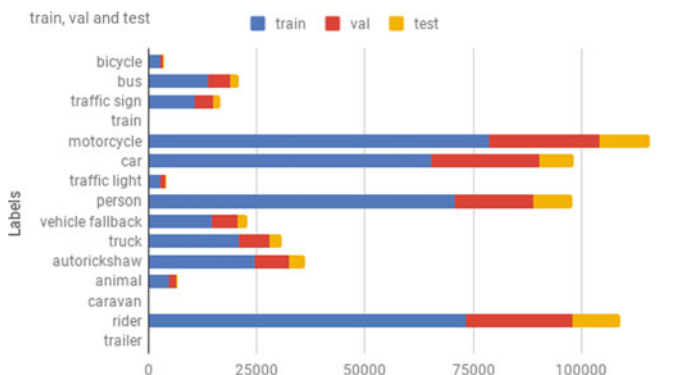
Data was collected from two sources—India Driving Dataset<sup>2</sup> by IIIT Hyderabad, and CVC FIR Dataset [4]. India Driving Dataset contains manually annotated bounding boxes for 15 distinct classes as shown in Fig. 1. The CVC FIR is also a dataset for object detection, which uses a far-infrared camera to capture images.

These classes were rearranged, and validation checks were performed to remove erroneous annotations. The classes ‘train’, ‘traffic light’, ‘animal’, ‘caravan’, ‘trailer’ were not required and hence were removed. The class ‘bicycle’ was renamed and merged into ‘motorcycle’, ‘vehicle fallback’ was merged into ‘car’, and ‘rider’ was merged into ‘person’. The resulting 7 classes that were used to train an object detector model are shown in Fig. 2. Image augmentation was performed through random cropping for increasing the dataset size.

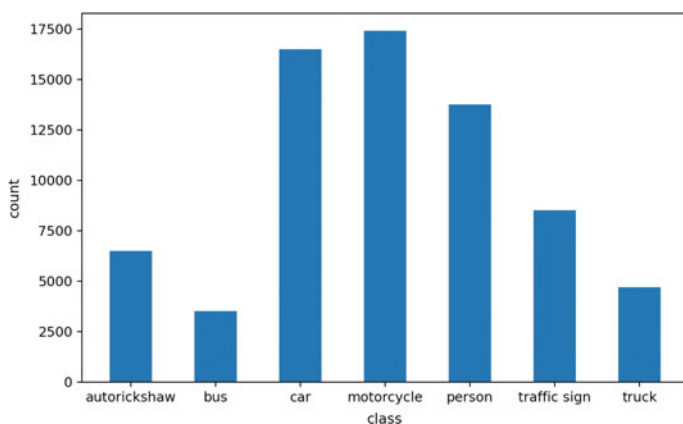
Out of these seven classes, ‘autorickshaw’, ‘bus’, ‘car’ and ‘motorcycle’ were considered as a ‘vehicle’ for further processing after the image goes through the object

---

<sup>2</sup> <https://idd.insaan.iiit.ac.in/>.



**Fig. 1** India Driving Dataset class distribution



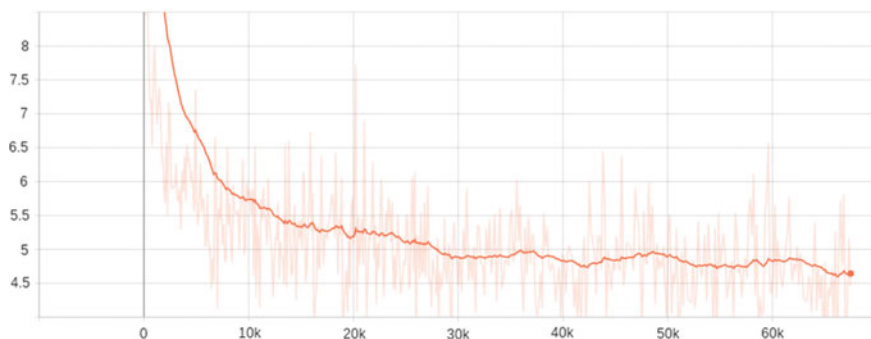
**Fig. 2** Distribution of the classes used for training

detector model. The CVC FIR dataset only contains annotations for pedestrians. Similar preprocessing steps were taken to create a usable dataset in the appropriate format.

## 4.2 Training

For both the datasets, an SSD MobileNet V2 [13] architecture was used. TensorFlow Object Detection API with Model Zoo<sup>3</sup> was used with models pretrained on the COCO dataset [9] as the base models. Feature extractor layers from the base models were left untouched but by using transfer learning, and the final dense layers were

<sup>3</sup> [https://github.com/tensorflow/models/tree/master/research/object\\_detection](https://github.com/tensorflow/models/tree/master/research/object_detection).



**Fig. 3** Training total loss

trained to localize an object in the image and also classify it accordingly. A batch size of 16 was used along with the Adam optimizer with an initial learning rate of 0.0002 [7]. The total loss while training for around 67 thousand epochs is shown in Fig. 3.

### 4.3 Traffic Sign Classifier Model

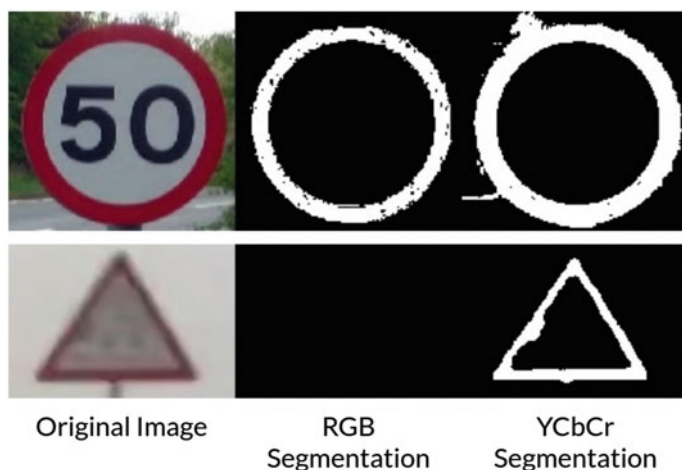
With the rise in the research for self-driving vehicles, the need for an efficient traffic sign detection system was inevitable; as stated by Wang et al. [16] The traffic sign classifier model aims to classify the traffic signs according to their shapes—‘circle’, ‘triangle’ and ‘other’. The motive behind doing this is that in the context of Indian traffic signs, each shape has some meaning associated with it. These are outlined in Indian Road Congress (IRC) 67:2012 (Code Practice for Road Signs).<sup>4</sup> Another reference is Code of Practice (Part 4) Signages<sup>5</sup> prepared by the Institute of Urban Transport and the Ministry of Urban Development, based on the original works of the IRC. In essence, circles depict regulatory signs, triangles depict warning signs, and rectangles depict informative signs.

A dataset of such traffic signs was prepared by extracting them from the India Driving Dataset. The object detection model was used to perform inference on those images along with videos scraped from YouTube to obtain bounding boxes of detected traffic signs. These were cropped and manually segregated into three categories—circle, triangle and other, generating a total of 1600 images. Image augmentation was performed through horizontal flips, slight rotation and zoom to increase the dataset size.

Since Indian traffic signs have a red boundary, colour segmentation was used to extract red colour from the images to identify the boundary. This creates a binary mask, where a pixel value of 1 represents red colour and 0 represents not red. The

<sup>4</sup> [https://archive.org/details/govlawircy2012sp67\\_0](https://archive.org/details/govlawircy2012sp67_0).

<sup>5</sup> <http://164.100.161.224/upload/uploadfiles/files/IUT-4.pdf>.

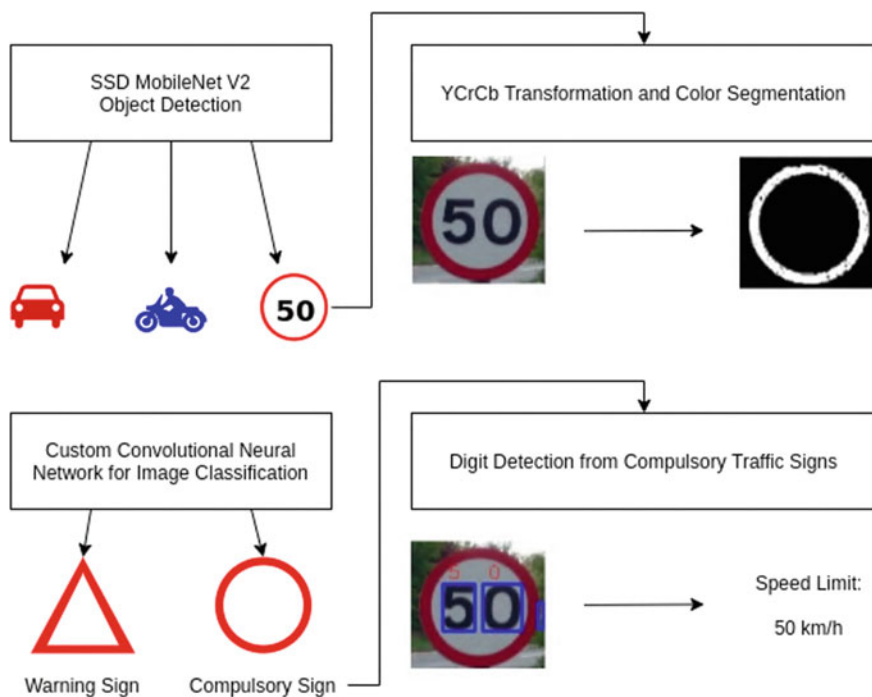


**Fig. 4** RGB vs YCbCr for segmentation

original image was a 3D matrix (height \* width \* 3 channels for RGB), whereas the resulting image is a 2D matrix with binary pixel values. This greatly reduces the complexity of the problem of classifying traffic signs into shapes. Since the model for this task is smaller and simpler, it allows for faster training and execution. This is extremely crucial as far as processing video in real time is concerned.

One of the biggest obstacles while extracting red colour from the images was variation in lighting conditions. RGB isn't a particularly good colour space for this purpose, since any colour filter is highly dependent on the overall light intensity. Different colour spaces were experimented with, and it was found that YCbCr gives substantially better results. YCbCr is a colour space that is often used in media formats like JPEG. It has been briefly described by Kolkur et al. [8]. The three channels that YCbCr uses are luma, blue-difference and red-difference, respectively. Because there is a separate channel for luminosity, it has great applications in media transmission, where luminosity and chromaticity can be compressed individually depending on the use case. For example, since the human eye is more susceptible to light intensity than colour, the luminosity channel can be assigned a higher bit-depth than the chromaticity channels. This allows optimum compression while retaining image quality for the human eye. It also helps the colour filter to exclude luminosity completely for colour segmentation. This way, the module can identify red colour irrespective of the lighting conditions in which the image is captured. The difference between RGB and YCbCr for two images is shown in Fig. 4. An RGB colour filter with one set of parameters works for one image, but does not generalize at all. This is not the case with a YCbCr filter.

The classification model is built using Keras library. It is a simple convolutional neural network with three convolutional layers at the top combined with pooling layers, followed by 1 dense layer with an output of 3 units.



**Fig. 5** Model pipeline

#### 4.4 Digit Recognizer Model

As surveyed by Islam et al. [6], optical character recognition is typically done by first identifying the location of the characters, followed by classifying it as one of the characters. This two-step approach often is at times slow to execute. In this scenario, performance is crucial, and hence, a better approach is to use an object detection model which will find the bounding box of the digit along with the digit in one step. The digit recognizer model from Penny 2018<sup>6</sup> was used to identify the digits on traffic signs. The speed limit was extracted by arranging the digits according to their  $x$ -coordinates. The entire resulting pipeline is demonstrated in Fig. 5.

#### 4.5 Calculating Driving Speed

The final calculation of the driving speed depends on the detected objects and their distance from the camera. Sadreddini et al. [12] present an approach using a single

<sup>6</sup> <https://github.com/penny4860/SVHN-deep-digit-detector>.

camera, but it is limited to indoor measurements as it relies on edge detection of walls. Chen et al. [2] present a method to estimate the distance of vehicles by exploiting the standard dimensions of number plates. This paper demonstrates the estimation of distance by finding the area of the bounding box that is drawn around the object. It allows the model to be used outdoors and also eliminates the dependency on standard object sizes, enabling this approach to be extended to pedestrians as well. Megalingam et al. [10] present a method to estimate a safe distance while tailgating a vehicle in traffic or while parking, but it does not provide a mechanism for calculating a safe speed limit with which a vehicle should proceed.

In optical physics, the size of an image and its distance from the lens exhibit an inverse linear relationship.<sup>7</sup> The size is a physical dimension, but this paper presents an approach to approximate it to digital pixels in an image depending on the camera orientation and its field of view (FoV). To start with, the height of the bounding box  $h$  can be multiplied by a scaling factor  $k$ , to produce the distance  $d$ . Sometimes, the orientation of the object makes the bounding box appear taller or wider than normal. In such cases, it is helpful to consider the width  $w$  of the box along with the height. However,  $h * w$  is no longer linearly proportional to the distance, and hence, a square root is taken.

$$d = \frac{k}{\sqrt{h * w}} \quad (1)$$

It is also wise to express  $h$  and  $w$  in terms of the percentage of height and width of the actual image, rather than the number of pixels. This eliminates its dependency on the resolution of the sensor on which the image was captured. Also, the values need not be changed after any resizing operations which are required while training a deep learning model. For example, an object of size  $40 \times 70$  px in an image of  $100 \times 100$  px resolution will have the dimensions represented as 0.4 and 0.7, respectively.

Finding a suitable value of  $k$  in Eq. 1 is dependent not only on the camera's FoV (its optical zoom) and its orientation but also on the relative size of an object in the camera's frame. For example, a car with pixel dimensions of  $100 \times 100$  px is going to be much farther than a coin with the same pixel dimensions, simply because a car is much larger than a coin.

To use this information in finding the appropriate driving speed, an impact value by each object is calculated. It represents the impact that it is going to have on the speed negatively; i.e. the higher the impact, the lower the recommended speed will be. A vehicle's impact or  $I_V$  is calculated as the inverse of its distance from the camera.

$$I_V = \frac{1}{d} \quad (2)$$

For pedestrians, an additional parameter is added to the impact—their closeness to the horizontal centre of the image. The motive behind this extra parameter is that pedestrians walking on the footpath or sidewalk are not in the way of the vehicle

<sup>7</sup> <https://www.khanacademy.org/science/physics/geometric-optics/lenses/v/object-image-height-and-distance-relationship>.

**Table 1** Comparison of mAP

Model	IoU	mAP
Faster RCNN	0.5	0.31
SSD MobileNet V2	0.5	0.36

and should not have much of an impact. However, pedestrians that are close to the horizontal centre are much more likely to be a problem and hence have a higher impact. This horizontal impact or  $I_h$  is calculated as the difference between 0.5 (centre of the image) and the horizontal deviation of the centre of the bounding box from the centre of the frame. Equation 3 compiles this in a formula, where  $c_x$  represents the  $x$ -coordinate of the centre of the object's bounding box.

$$I_h = \frac{1}{d} * (0.5 - |c_x - 0.5|) \quad (3)$$

The appearance of the extra term  $\frac{1}{d}$  in Eq. 3 is explained by the fact that  $I_h$  also depends on the distance of the pedestrian since a pedestrian far away with a lower horizontal deviation is still likely to be on the sidewalk. This is just like how railway tracks in a photograph seem to converge in the distance, but in reality, they always stay parallel to each other. The pedestrian's impact or  $I_p$  is calculated as shown in Eq. 4.

$$I_p = \frac{1}{d} + I_h \quad (4)$$

The final speed to be recommended or  $V$  is calculated from an initial speed  $U$ .  $U$  is chosen dynamically based on external factors viz. type of road, traffic signs, proximity to schools and hospitals in urban areas using the Nominatim API.  $I_v$  and  $I_p$  are penalties that are applied to  $U$ . The impact for all pedestrians and vehicles is summed, multiplied with a weight ( $W_v$  for vehicles and  $W_p$  for pedestrians) and subtracted from the initial speed  $U$ . The resulting formula for  $V$  is shown in Eq. 5.

$$V = U - W_v \sum I_v - W_p \sum I_p \quad (5)$$

## 5 Results

Bhargava [1] describes the results of training Faster RCNN on the India Driving Dataset. A comparison with SSD MobileNet V2 proposed in this paper is shown in Table 1. This model achieves significantly better results while being lighter in size than faster RCNN by design.

**Fig. 6** Speed sign identification



**Fig. 7** Warning sign identification



**Fig. 8** Crowded road



The colour segmentation process with the image classification model together achieves a categorical accuracy of 92.46%

Some snapshots of the complete pipeline working in real-life conditions are shown in Figs. 6, 7, 8 and 9. The number at the top left corner represents the final recommended speed for the vehicle. These frames are from videos that were scraped from YouTube.<sup>89</sup>

<sup>8</sup> <https://www.youtube.com/watch?v=0ZH0-R3mrtQ>.

<sup>9</sup> <https://www.youtube.com/watch?v=KWJaBJYJIjI>.



**Fig. 9** Detections on CVC IR dataset



## 6 Conclusion

As per our thorough research, this solution is economically viable—its cost is almost negligible as compared to the total price of an average automobile in India. The system uses just one camera at the moment, which can hamper the accuracy in estimating the distance. In future, the accuracy of the system could be increased by 3D object detection and sensor fusion using two cameras, lidar, radar and other sensors. The module uses edge computing and works perfectly in areas where there is low Internet connectivity. This feature helps to cater to large audiences, as Internet connectivity might be an issue in some areas. Due to the unique proposition of this solution, the module can be also used in low light conditions. Thus, due to the novelty of this solution, the system is able to accurately notify the safe speed limit to the driver even in extreme conditions. This will ensure the safety of the driver as well as pedestrians, thus reducing the number of road accidents by a significant amount. It will particularly thrive in areas with lot of traffic and pedestrians, such as prime locations in cities. To summarize, by adapting our cutting-edge technology in the traditional automobile industry, we can hope to have smart cities where we will have accident-free neighbourhoods.

## References

1. Bhargava P (2019) On generalizing detection models for unconstrained environments. In: 2019 IEEE/CVF international conference on computer vision workshop (ICCVW) . <https://doi.org/10.1109/iccvw.2019.00529>. <http://dx.doi.org/10.1109/ICCVW.2019.00529>
2. Chen S-H, Chen R-S (2011) Vision-based distance estimation for multiple vehicles using single optical camera. In: 2011 2nd international conference on innovations in bio-inspired computing and applications. 2011, pp 9-12. <https://doi.org/10.1109/IBICA.2011.7>.

3. Frank R, Forster M (2014) A recommendation based driver assistance system to mitigate vehicular traffic shock waves. In: IEEE vehicular networking conference (VNC). IEEE (2014), pp 125–126
4. Gonzalez AA et al (2016) Pedestrian detection at day/night time with visible and FIR cameras: a comparison. *Sensors* 16:820. <https://doi.org/10.3390/s16060820>
5. Hosseinlou MH, Kheyraadi SA, Zolfaghari A (2015) Determining optimal speed limits in traffic networks. In: *IATSS Res* 39(1):36–41. ISSN: 0386-1112. doi: <https://doi.org/10.1016/j.iatssr.2014.08.003>. <https://www.sciencedirect.com/science/article/pii/S0386111214000259>
6. Islam N, Islam Z, Noor N (2017) A survey on optical character recognition system. 2017. [arXiv: 1710.05703](https://arxiv.org/abs/1710.05703) [cs.CV]
7. Kingma DP, Adam JB (2017) A method for stochastic optimization. [arXiv: 1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG]
8. Kolkur S et al (2017) Human skin detection using RGB, HSV and YCbCr color models. In: *Proceedings of the international conference on communication and signal processing 2016 (ICCASP 2016)*. <https://doi.org/10.2991/iccasp-16.2017.51>.
9. Lin T-Y et al (2014) Microsoft COCO: common objects in context. In: David F et al (eds) *Computer vision ECCV 2014*. Springer, Cham, pp 740–755. ISBN: 978-3-319-10602-1
10. Megalingam RK et al (2016) Monocular distance estimation using pinhole camera approximation to avoid vehicle crash and back-over accidents. In: *2016 10th international conference on intelligent systems and control (ISCO)*, pp 1–5. <https://doi.org/10.1109/ISCO.2016.7727017>
11. Pathan KH, Patil MM (2017) Recommendation based advance driver assistance system: using Raspberry Pi. In: *2017 international conference on intelligent computing, instrumentation and control technologies (ICICICT)*. IEEE 2017, pp 1240–1247
12. Sadreddini Z, Cavdar T, Jond HB (2016) A distance measurement method using single camera for indoor environments. In: *2016 39th international conference on telecommunications and signal processing (TSP)*, pp 462–465. <https://doi.org/10.1109/TSP.2016.7760921>
13. Mark S et al (2019) MobileNetV2: inverted residuals and linear bottlenecks. [arXiv: 1801.04381](https://arxiv.org/abs/1801.04381) [cs.CV]
14. Sundar R, Hebbar S, Golla V (2014) Implementing intelligent traffic control system for congestion control, ambulance clearance, and stolen vehicle detection. *IEEE Sens J* 15(2):1109–1113
15. Várhelyi A (2002) Dynamic speed adaptation in adverse conditions: a system proposal. *IATSS Res* 26(2):52–59. ISSN: 0386-1112. [https://doi.org/10.1016/S0386-1112\(14\)60043-1](https://doi.org/10.1016/S0386-1112(14)60043-1). <https://www.sciencedirect.com/science/article/pii/S0386111214600431>
16. Wang C (2018) Research and application of traffic sign detection and recognition based on deep learning. In: *2018 international conference on robots intelligent system (ICRIS)*, 2018, pp.150–152. <https://doi.org/10.1109/ICRIS.2018.00047>

# Real-Time Markerless Facial Landmark Detection Using Deep Learning



Samyak Agarkar  and Kapil Hande 

**Abstract** Locating facial landmarks like eyes, eyebrows, nose, lips, facial outline, etc. can be used as the foundation for generating facial deformations caused by expressions and head movements, which is very crucial in 3D animation, the VFX industry, the gaming industry, VR & AR. This project intends to focus on building a robust and real-time Facial Landmark detection model by improving upon existing models. Over the past few years, a lot of research has been done for detecting facial landmarks without the use of markers. Yet, there is a huge scope for improvement in terms of accuracy and speed. There are three major types of facial Landmark detection algorithms: Holistic methods, Constrained Local Model (CLM), and regression-based methods. Holistic methods try to build a Global representation of facial appearance accurately but require a lot of processing power and time. CLM Build a local representation of facial appearance whereas, regression-based methods simply capture the facial shape and appearance information. Classical regression techniques are fast but have many limitations, for instance, they fail to detect landmarks in occluded faces. Neural networks and deep learning have opened up new possibilities in building models that are efficient, robust, and dependable. Though deep learning models are processor-intensive tasks, recent improvements in network compression and pruning have made it possible to use deep learning models in real-time on fairly low specification computers. This study intends to examine new combinations of previous implementations, new ideas, and heuristic information. The proposed model will be better in terms of speed than other models and on par if not better in terms of accuracy.

**Keywords** Facial landmark detection · CNN · Deep learning · Machine learning · AI

---

S. Agarkar (✉) · K. Hande

CSE Department, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

## 1 Introduction

### *1.1 Need of Facial Landmark Detection*

Facial expressions are one of the most important aspects of Visual communications in humans as well as animals. In this digital era, there are many fields where we need to replicate this form of visual communication. Fields like 3D animation, the VFX industry, game industry, audiovisual learning, VR, and AR require facial expressions to be generated wherever human or animal is involved. Traditionally, Facial expression capturing or facial motion capture has been done by tracking markers on the face of motion capture artists to drive shape keys of 3D characters designed in computers. The equipment required for such a setup is very costly and only big studios can afford it. This becomes a hurdle for small studios and individuals in using Facial Expressions in short movies and renders. Advancements in computer vision made it possible to detect facial landmarks without the use of markers. This made it cheaper and easy for anyone to drive facial expressions of 3D characters using video. The recent growth of the VR and AR field led to the development of good facial landmark detection algorithms which are fast enough to work on handheld mobile phones. Yet, the Landmark detection techniques have seen a tradeoff between accuracy and speed. AR filters used in consumer digital products like Snapchat and Instagram use facial Landmark detection at its Core. While using these filters it is evident that, though the algorithm works very fast, its accuracy suffers a huge compromise. A lot of research is being done in this field to improve accuracy as well as speed. Still, there is a lot of scope for improvement.

### *1.2 Deep Learning as a Solution*

Deep learning methods required a huge data set to train and test the neural networks. For facial landmark detection, there are various datasets available. These datasets have been built from images in the wild and then manually annotated to extract the landmark features. Facial landmark detection is a two-step process. First, we need to identify or detect a face in an image. This can be easily done using existing face detection models like Haar-Cascade, Voila-Jones, or using face detection by CNN. The next step is to detect the landmarks on the face. Landmarks need to be detected in such a way that sufficient information is available to closely if not accurately reconstruct the exact facial expression. Various available data sets cater to various numbers of landmarks. MTFI takes 5 points into consideration, AFW uses 6 points, AFLW uses 21 points and 300 W-LP uses 68 feature points for facial landmarks. Definitely output and its accuracy will differ for each dataset.

An image is just a set of pixels and by itself cannot be provided as input to a neural network as without filtering they possess no useful information. Convolutional networks use convolutions or filters to extract useful information from the image.

Filters like edge filters, contour filters, etc. are applied to the images to extract useful information before feeding into the neural network. Hence a convolutional neural network consists of convolutional filters followed by neural network dense mesh including hidden layers as well as output layer.

## 2 Literature Survey

### 2.1 *Classical Methods*

The first attempts at Facial Landmark detection were the use of template fitting methods. These techniques tried to fit a face template onto the given image. Limitations could be seen when face orientations cause problems in fitting the template. Active shape models e.g. [1, 2] tried to fit the template of a basic shape on the image and extract locations off the feature points. The next set of techniques involved regression-based models which tried to detect the landmarks using regression methods. e.g. [3, 4]. Such techniques are still used wherever accuracy can be relaxed. These models take care of facial orientations but there are chances of landmark overlapping and facial occlusions are not taken care of.

### 2.2 *Deep Learning Methods*

The development of deep learning and the use of convolutional neural networks showed a new way to extract facial landmark information using deep learning techniques. Convolutional neural networks or CNNs can extract useful information from images and learn to reproduce them under a supervised environment. There are a few unsupervised learning methods as well but controlling them becomes difficult. Dong et al. [5] uses style-aggregated face generation models. Chen et al. [6] uses structured predictions using hyper-parameters and heuristics. A few existing CNN methods have also used Hierarchical dividers on cascaded convolutionals [7]. Attention-driven cropping of faces [8] from images has shown good improvements in removing feature overlap. Recursively the images are divided until each leaf node is a feature. This leads to a neural network to learn effectively. Reducing the image into a subset of features and then expanding again to learn from those feature points has proven to be effective. But this also leads to more hidden layers and hence more computation time during training as well as testing and real-world usage.

Here is a list of few notable methods used in academic research environments (Table 1).

**Table 1** Notable methods and their citations

Method	Description	Paper
TCDCN	Tasks-constrained deep model, with task-wise early stopping to facilitate learning convergence	Zhang et al. [9]
HyperFace	Fuses the intermediate layers of a deep CNN using a separate CNN followed by a multi-task learning algorithm that operates on the fused features	Ranjan et al. 2016
Consensus of exemplars	Combines the output of local detectors with a non-parametric set of global models for the part locations based on over one thousand hand-labeled exemplar images	Belhumeur et al. [10]
3DDFA	A dense 3D face model is fitted to the image via a convolutional neural network (CNN)	Zhu et al. [11]
CFAN	The Cascades a few successive Stacked Auto-encoder Networks (SANs)	Zhang et al. [9]
CFSS	Stage-by-stage progressive and adaptive search of the face instead of fitting a predefined template	Zhu et al. [12]
SDM	Learns a sequence of descent directions that minimizes the mean of NLS functions sampled at different points	Xiong and De la Torre Frade [13]
3D regression	Based on a cascade regression framework that directly estimates face landmarks locations in 3D	Tulyakov and Sebe [14]
Explicit shape regression	Directly learn a vectorial regression function to infer the whole facial shape from the image and explicitly minimize the alignment errors over the training data	Cao et al. [15]

**Table 2** List of datasets with landmark points

Dataset name	Landmark points
MTFL	5
AFW	6
AFLW	21
300 W-LP	68

## 3 Methodology

### 3.1 Selection of Dataset

There are many widely used in-the-wild dataset for faces which can be used to train our model. The number of features all landmarks that are noted in the data sets vary widely. That we will be using to train and test the system (Table 2).

300 W-LP Dataset Maximum Landmark points should give most accurate results. One notable drawback of having more landmark points is more processing would be required, hence, slowing down the system. Thus, the tradeoff between number of landmark points and runtime will be crucial.

### 3.2 Tools Used

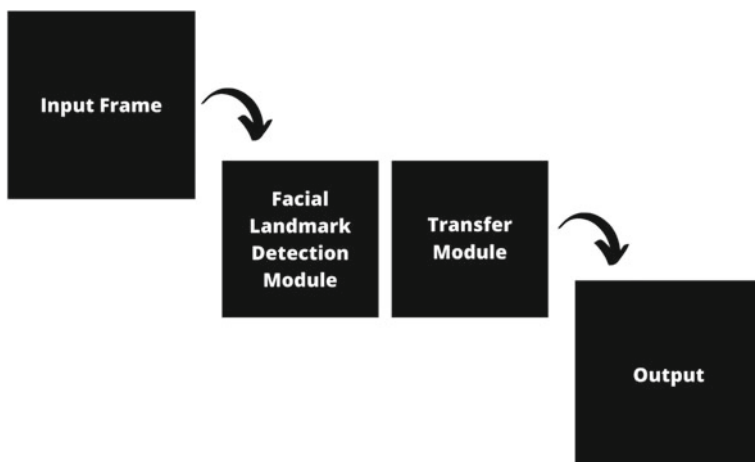
#### Tensorflow and Keras

Tensorflow is a free and open source software library developed by Google, used for data flow and differentiable programming across a large range of tasks. This library takes advantage of multiple CPUs and GPUs. It has built in support for parallel processing and clustered processing. This makes tensorflow a very good choice for deep learning algorithm processing.

Keras is an open source library that provides a Python interface for artificial neural networks. It acts as an interface for the tensorflow library. Keras supports standard neural networks as well as convolutional and recurrent neural networks. The biggest advantage of using keras is that we don't need to worry about the data flow between the layers. The convolutional neural network or CNN the has been built using Tensorflow and Keras.

#### OpenCV

OpenCV is an open source computer vision library which has many image editing and computer vision based algorithms built into it. The system uses OpenCV's face detection algorithm (Fisherface) to detect the faces in the frame. Some image transformations are also done using OpenCV library.



**Fig. 1** System flow diagram

### 3.3 The Proposed System

The proposed system has two main modules (See Fig. 1).

**Facial Landmark detection module** This module detects a face in the frame and outputs the relative coordinates of landmarks detected on the face. It has two sub-modules.

*Face detection.* This sub-module finds a face in the frame so that the input data for the next module is lowered by cropping in the face. Fisherface algorithm has been used from OpenCV library to efficiently detect faces from a frame.

*Landmark detection.* This sub-module is convolutional neural network (CNN) with some hierarchical cascade division using hyper parameters to improve the efficiency of the system. This module is the main focus of the research project. There is, of course, a tradeoff between run-time and accuracy for the model. This tradeoff is important to note as we are trying here to get a real-time facial capture. Details about the CNN model used are discussed in the next section.

**Transfer Module.** The transfer module transfers the landmark detection output to a 3D human face model so that we can test out our system in real life applications. The transfer of data is done by mapping the first module's output to the input of a 3D application (Blender) using either a pre-made or custom-made driver.



### 3.4 The CNN Model

#### The Model

The CNN sequential model built using tensorflow and keras has the following specifications.

- OpenCV face detection algorithm detects face in frame. The face image is then cropped and transformed back to  $450 \times 450$  px image.
- The input layer is  $450 \times 450$  px divided into 3 channels for RGB. This is in accordance with the image data set image size.
- The input layer is followed by a sequence of 5 Max pooling layers and convolutional 2D layers in an alternate fashion. These layers are intended to extract information from the Frame such as lines, feature points extra. We use ReLU activation function for each convolutional layer.

ReLU:

```
if input > 0: return input.
else: return 0.
```

- The flatten layer flattens the convolutional Matrix output to one single array that acts as input to the dense layer
- The pen-ultimate dense layer gives output of 256 numbers and activation function of ReLU
- The final dense layer outputs 136 floating point numbers. These 136 floating point numbers corresponding to  $X$  and  $Y$  coordinates of 68 Landmark points as in the data set.  $68 \times 2 = 136$
- We use Adam optimizer and mean squared error as a loss function to the model. Accuracy is used as a metric that will be eventually used in backpropagation.

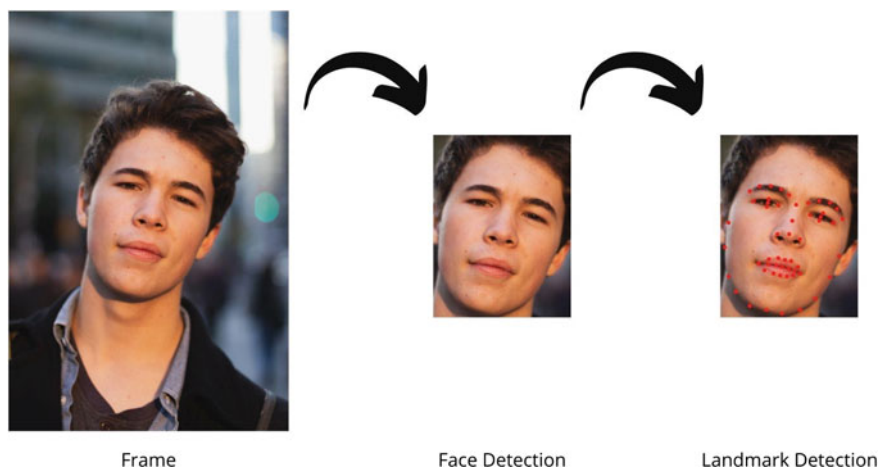
Cropping the frame to accommodate on the face makes it much easier for the CNN model to locate features on the frame. Convolutional 2D layers generate random filters and apply them on each frame during a layer transit. Characteristics like edge, contour, lines, hue change, etc. are expected to be found out by these layers. It is assumed that convolutional layers learn to separate out features of the face during the trainin. The flatten layer flattens the 2D output into and 1D array which is then be inputted to a dense neural network layers which then extracts the exact locations of the required facial features. During the training of the model, all the hyper parameters are initially randomly generated. Iteration over each image produces some output which is then compared to the expected output during the supervised learning. The error is them back propagated from the output layer back to the input layer which changes the bias and parametric values of inputs of each node in each layer of the neural network. Iterating over the entire 65,000 images is done in batches. Accuracy of the model after each iteration has been noted in the experimental results section.

## 4 Experimental Results

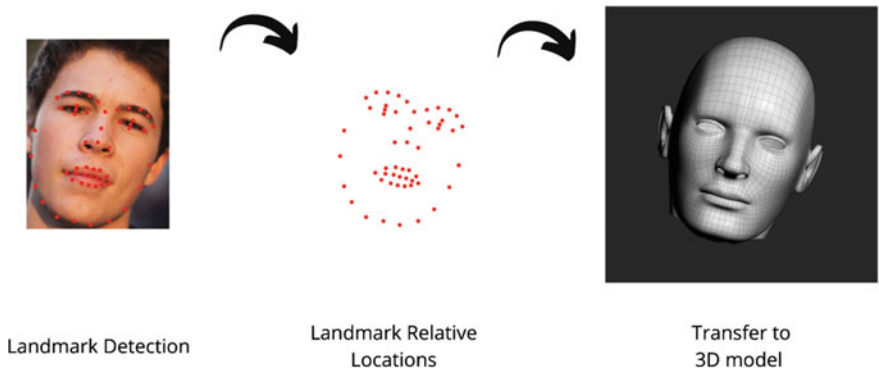
The model was trained on a personal laptop as well as on Google colab. Here is a list of accuracy at each epoch that was performed on the model.

As can be seen from the Fig. 4, initially the training is faster. As the model runs for more number of epochs, accuracy increases in a logarithmic fashion. Each epoch took around 8 h on the laptop (NVIDIA 940MX GPU) to complete. Following the graph we can safely assume that the sequential model we built is working and the accuracy will tend to 100% but never reach its full potential. It should also be noted that there is very little chance of overfitting as the number of epochs is very less as compared to the number of samples that we have in the data set. We can safely assume that the model sees a different image every time and does not overfit on any single image due to suffering applied during the training process in between each epoch.

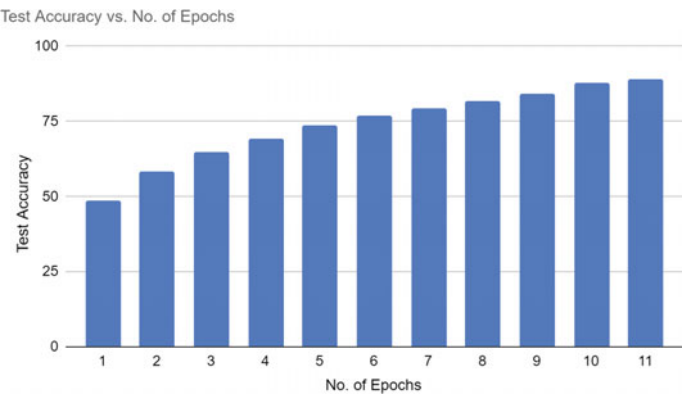
Compared to the previous systems as discussed in the literature survey section. The current model stands in the middle ground both on accuracy as well as speed terms. The current model is a little bit faster then more accurate models like Hyperface and 3DDFA, but it is little less inaccurate than them. The model is slower than a few faster models like SDM but also tends to be more accurate then SDM. As discussed earlier there is a definite trade off between speed and accuracy. The current model should hence stand as a middle ground between a models which are more accurate but slower and models which are more faster but inaccurate. Figures 2 and 3 are the outputs generated by the system when transferring the output of CNN to blender 3D face model. As can be seen by the figures, the landmarks detected by the system are very closely in accordance to the real features of the face. OSC live feed of detected landmarks to Blender model gave around 25 FPS. It should be noted that



**Fig. 2** Module 1: Facial Landmark detection module



**Fig. 3** Module 2: Transfer module



**Fig. 4** Testing accuracy versus no. of epochs

slow performance of the model is a constraint because of the system processing power as discussed in the limitations section. Figure 4 shows the test accuracy at the recorded epochs.

## 5 Conclusion

There is a huge demand of facial feature detection systems in VFX and animation industries. The proposed model take the middle ground in the trade off between speed and accuracy. With proper training and integrations of the model to other softwares, the proposed model can be efficiently used to drive facial animations of human as well as animal characters. This study can also be used as the basis for more more

pronounced deep learning applications like face recognition and face morphing. More study can still be done to close the gap between the speed-accuracy trade off.

## 6 Limitations of the Study

Training Period of the CNN model was very constant due to time limitations. It is expected that the accuracy would increase with more training iterations. comparatively less powerful GPU and the system was used for or the entire study. Using more powerful industry-standard GPUs like Nvidia Quadro or RTX shall speed up the process and also give much faster results during training as well as testing. Better datasets and more clear images using good webcams should also help the model train much properly and give more accurate results.

## References

1. Ahlberg J (2002) An active model for facial feature tracking. *EURASIP J Adv Signal Process* **2002**:569028. <https://doi.org/10.1155/S1110865702203078>
2. Alabort-i-Medina J, Zafeiriou S (2014) Bayesian active appearance models. In: 2014 IEEE conference on computer vision and pattern recognition, pp 3438–3445. <https://doi.org/10.1109/CVPR.2014.439>
3. Cristinacce D, Cootes T, Boosted regression active shape models. In: British machine vision conference, pp 880–889. <https://doi.org/10.5244/C.21.79>
4. Dantone M, Gall J, Fanelli G, Van Gool L (2012) Real-time facial feature detection using conditional regression forests. In: 2012 IEEE conference on computer vision and pattern recognition, pp 2578–2585. <https://doi.org/10.1109/CVPR.2012.6247976>
5. Dong X, Yan Y, Ouyang W, Yang Y (2018) Style aggregated network for facial landmark detection. In: 2018 IEEE/CVF conference on computer vision and pattern recognition, pp 379–388. <https://doi.org/10.1109/CVPR.2018.00047>
6. Chen L, Su H, Ji Q (2020) Deep structured prediction for facial landmark detection. [arXiv: 2010.09035](https://arxiv.org/abs/2010.09035) [cs.CV]
7. Zhu S, Li C, Loy C-C, Tang X (2016) Unconstrained face alignment via cascaded compositional learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3409–3417
8. Chandran P, Bradley D, Gross M, Beeler T (2020) Attention-driven cropping for very high resolution facial landmark detection. In: 2020 IEEE/CVF conference on computer vision and pattern recognition (CVPR), 2020, pp 5860–5869. <https://doi.org/10.1109/CVPR42600.2020.00590>
9. Zhang J, Shan S, Kan M, Chen X (2014) Coarse-to-fine auto-encoder networks (CFAN) for real-time face alignment. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) *Computer vision—ECCV 2014*. ECCV 2014. Lecture notes in computer science, vol 8690. Springer, Cham. [https://doi.org/10.1007/978-3-319-10605-2\\_1](https://doi.org/10.1007/978-3-319-10605-2_1)
10. Belhumeur PN, Jacobs DW, Kriegman DJ, Kumar N (2013) Localizing parts of faces using a consensus of exemplars. *IEEE Trans Pattern Anal Mach Intell* **35**(12):2930–2940. <https://doi.org/10.1109/TPAMI.2013.23>
11. Zhu X, Lei Z, Liu X, Shi H, Li SZ (2016) Face alignment across large poses: A 3D solution In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp 146–155

12. Zhu S, Li C, Change Loy C, Tang X (2015) Face alignment by coarse-to-fine shape searching. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp 4998–5006
13. Xiong X, De la Torre F (2013) Supervised descent method and its applications to face alignment. In: Conference paper, proceedings of (CVPR) computer vision and pattern recognition, pp 532–539
14. Tulyakov S, Sebe N (2015) Regressing a 3D face shape from a single image. In: 2015 IEEE international conference on computer vision (ICCV), pp 3748–3755. <https://doi.org/10.1109/ICCV.2015.427>
15. Cao X, Wei Y, Wen F, Sun J. Face alignment by explicit shape regression. In: Proceedings/CVPR. IEEE Comput Soc Conf Comput Vision Pattern Recog 107:2887–2894. <https://doi.org/10.1109/CVPR.2012.6248015>

# A Novel Approach for Forecasting Account Receivables



Parth Kapadia, Bintu Kadhiwala, Tejaswini Bahurupi, Het Dalal, Siddhi Jariwala, and Kshitij Naik

**Abstract** In recent times, various firms/companies provide services to their customers on credit. These firms receive the payments, against the services and/or purchases provided to the customers but not paid for, termed as Account Receivables. With an aim to plan the finance in addition to decide strategies for their business, the executives of these firms desire to predict the Account Receivables of their companies for a specific time span. For the said purpose, in this paper, we present three different methods. These methods utilize machine learning prediction models—logistic regression, random forest classifier, and an ensemble of  $K$ -Means clustering and random forest classifier discussed in the literature. However, with an aim to find the optimal solution in terms of prediction accuracy, we put forward the theoretical and the experimental analysis of these methods. Experimental results affirm that an ensemble of  $K$ -Means clustering and random forest classifier outperforms among three methods.

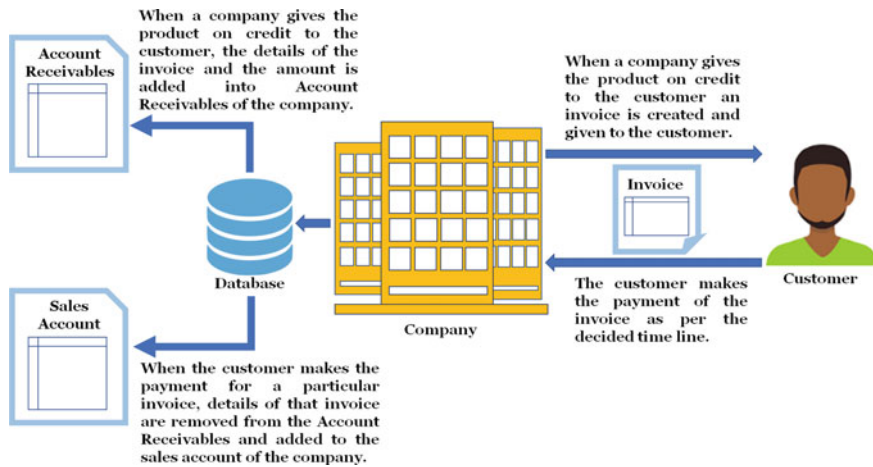
**Keywords** Account Receivables · Logistic regression · Random forest classifier ·  $K$ -Means clustering · Ensemble · Forecasting

## 1 Introduction

In the corporate world, most of the customers purchase the goods/services on credit and promise to pay the amount to the firm/company on a particular scheduled date or within a specific time span (Fig. 1). For the services and/or purchases provided to the customers but not paid for, the firms receive the payments from the customers popularly termed as ‘Account Receivables’. Usually, the companies give their customers a limit of 30, 60 or 90 day to pay and record this transaction as an invoice. The company lists this amount to be paid by the customer in its Account Receivables. Once the company receives the payment for the particular invoice, the amount is inserted into the sales account of the business.

---

P. Kapadia (✉) · B. Kadhiwala · T. Bahurupi · H. Dalal · S. Jariwala · K. Naik  
Computer Engineering Department, Sarvajani College of Engineering and Technology,  
Surat, India



**Fig. 1** Working of account receivables

Thus, Account Receivables is the money or the payment a customer owes to the company for the purchases they make on credit. On the balance sheet of the company, it is considered to be an asset for the company as it signifies the money that is to be received by the company. Therefore, Account Receivables is an important aspect for calculating the profits of the company. The executives of many companies allow a certain percentage of their sales to be on credit with an aim to increase their sales and to reduce the transaction cost [1]. As discussed in [2], Account Receivables exceed a quarter of the total corporate assets in countries like France, Germany and Italy. Moreover, the authors of [3] find that 18% of the total assets of US firms consist of receivables. As per our observations, the company must collect the outstanding amount within a certain time period to gain maximum benefits from Account Receivables.

The average number of days a company takes to collect Account Receivables amount from their customers is termed as 'Days Sales Outstanding' that gives an idea of how quickly a company can collect its outstanding money. The greater value for Days Sales Outstanding indicates that the company takes more time to collect money that can lead to cash flow problems. On the other hand, the lower value for the same indicates that the company collects the money in less time. In addition, the effectiveness of the company to convert the Account Receivables into cash is measured through the 'Receivables Turnover Ratio' that is a comparison between credit sales and average receivables [4]. It indicates how efficiently a company can collect the amount against the services provided to the customers on credit. Whenever the company fails to recollect this amount, it turns out to be a bad debt resulting in a loss for the company. Hence, with an aim to avoid such losses, it is highly required by the executives of the companies (1) to manage the Account Receivables, and (2) to forecast the Account Receivables expected to receive in a particular time span.

Consequently, forecasting of the Account Receivables plays a major role for the company executives as it can have a large impact on the working capital of the company. By forecasting, the company gets an insight into the cash amount to be

received from the customers against services over a certain time period. For this purpose, we train a model that predicts the probability of the customer paying the particular amount within 30, 60 and 90 day after the due date. Hence, the key objective of this paper is to find the optimal solution with the maximum prediction accuracy for forecasting Account Receivables. This forecasting definitely helps the company executives plan the finances of the company efficiently.

The rest of the paper is organized as follows: Sect. 2 discusses the related work. Section 3 describes various machine learning methods that we utilize for the purpose of the prediction and the real-time transaction dataset on which the experiments are performed. The experimental results obtained are debated and analysed at length in Sect. 4. Finally, Sect. 5 puts concluding remarks on our work.

## 2 Related Work

In [5], the authors work on the prototype that is able to support collectors in predicting the payment of invoices. They focus on the problems faced by the collectors to contact the customers on the payment due date for the collection of Account Receivables. For the same, they provide the solution through a predictive modelling approach to prioritize contacting clients based on the probability of late payment helps the collector to make a more effective and efficient decision. The major objective of their study is to make an assertive and timely decision by focusing on prioritizing the collection of invoices from the customers. This is decided based upon, (1) the possibility of higher financial return, and (2) the customers that are most likely to make the payment. They utilize different machine learning models such as naive Bayes, logistic regression, K-nearest neighbours, gradient boost decision tree, deep neural network. Their model is limited to prioritize the invoice payment and cannot predict when a particular payment will be made.

In [6], with the help of historical data, the author demonstrates how machine learning models predict payment outcomes of invoices that are not paid. Here, the author divides the days past due date into six buckets, viz. 1–30 days, 31–60 days and so on, in which the invoices can be paid and tries to predict the bucket in which the Account Receivables can be received. For the same, a predictive analysis approach is used to quantify the effect of the future decision in order to advise on possible outcomes before making the actual decision. Here, the author uses the historical data of the customers' paid invoices and tries to identify the payment behaviour of customers to create a predictive model. This model predicts the payment bucket of the newly credited invoice with the help of the AdaBoost decision tree algorithm. The accuracy attained by this approach is better than the approach discussed in [4]. However, the accuracy decreases gradually from bucket 1 to bucket 6 that ultimately results in only 50% accuracy for bucket 6. As described by the author in [6], the model is unable to handle those invoices which are overdue and cross the dispute limit.

In [7], the authors demonstrate how supervised learning is used to build predictive models for the payment outcomes prediction of newly created invoices. Their



model predicts whether an invoice is paid on time or not and also provides the delay estimation. Here, the task of prediction is formulated with the help of a model that classifies the new instance(s)/invoice(s) into five target classes. The authors study various algorithms, namely C4.5 decision tree induction, naive Bayes and PART algorithm and conclude that the PART algorithm gives the best performance in terms of classification accuracy. For the same, they develop a set of aggregated features that capture historical payment behaviour for each customer to make the prediction. Through this model, the company gets only a probable idea of whether the Account Receivables is paid by the customer or not. However, this model cannot predict the exact time span for paying the Account Receivables.

### **3 Theoretical Background**

#### ***3.1 Machine Learning Methods Considered***

In this subsection, we summarize three methods that we utilize for the purpose of the classification.

##### **3.1.1 Logistic Regression**

The examination of dichotomous or parallel classes is done with the help of a logistic regression model. Although it provides us with the capacity to adjust various classes, in our case we utilize the same to handle only two classes [8]. This makes strategic relapse particularly valuable for the investigation of observational information when change is expected to diminish. The normal method for fitting the logistic regression model has optimal properties [9]; however, we observe while performing experiments that it is extremely sensitive to noisy data. For experimentations, we build a logistic regression model in python using ‘sklearn.linear\_model.LogisticRegression’ package in scikit learn module.

##### **3.1.2 Random Forest Classifier**

The random forest classifier is an ensemble of several decision tree models for a set of data. For each tree, the data is split into units which are commonly referred as nodes. Split focus depends on estimations of indicator factors. Subsequently, factors used to part the information is viewed as significant illustrative factors. Random forest fits these separate decision trees to bring out a collective decision [10]. For experimentations and to have uniformity in findings, we also build the Random Forest classifier model in python using ‘sklearn.ensemble.RandomForestClassifier’ package in scikit learn module and we keep the ensemble of 100 decision trees.

### 3.1.3 Ensemble of K-Means Clustering and Random Forest Classifier

The use of ensemble classifiers is escalated recently as the main idea of the ensemble is to combine a set of models, each of which performs its own typical task with an aim to gain a composite model with improved accuracy. Alternatively, an ensemble classifier can also be viewed as a technique of combining many weak learners to produce a strong learner [11]. Additionally, in the past few years, experimental studies show that combining the outputs of multiple classifiers; i.e. ensemble methods reduce the generalization error [12]. In our work, we utilize an ensemble of the *K*-Means clustering and random forest classification algorithm as random forest classifier is not able to remove redundancies [13] that can be removed with the help of a clustering algorithm. As pre-processing is a process of removing redundant features, we employ a clustering approach using the *K*-Means algorithm to group similar features [14]. For experimentations, we pre-process the data using ‘sklearn.cluster.KMeans’ package and build the random forest classifier model in python using ‘sklearn.ensemble.RandomForestClassifier’ package in scikit learn module and we also keep the ensemble of 100 decision trees.

## 3.2 Transaction Dataset Description

We perform experiments on the live dataset provided by Sailfin Technologies Pvt. Ltd. [15]. The dataset consists of 51,53,556 closed transactions of one of the company’s customers. The dataset is comprising of the following six parameters:

1. ID—Transaction ID
2. Account Number—Account from which transaction takes place
3. Amount—Amount of the Account Receivables
4. Create Date—Date on which the Account Receivables is created
5. Due Date—Due date for closing the Account Receivables
6. Close Date—Date on which the Account Receivables is closed; i.e. full payment is made

## 3.3 Extracted Parameters

For experimentation, we build a classification model with the help of four extracted parameters. In this subsection, we summarize these parameters extracted from the transaction dataset (Sect. 3.2) along with the extraction methods.

1. Parameter—1 (Normalized Amount) and Parameter—2 (Normalized DPD): Using normalizing techniques over the datasets, the normalized amount and normalized DPD are calculated as follows:

- For method 1 (Sect. 3.1.1) and method 2 (Sect. 3.1.2):
  - (a) Normalized Amount = Actual amount/Average amount of dataset
  - (b) Normalized DPD = Actual DPD/Average DPD of dataset
- For method 3 (Sect. 3.1.3)
  - (a) Normalized Amount = Actual amount/Average amount of cluster
  - (b) Normalized DPD = Actual DPD/Average DPD of cluster

Where, Days Past Due (DPD) is calculated using,  $DPD = (\text{Due date} - \text{Close date})$

## 2. Parameter—3 (EPID) and Parameter—4 (EPDD):

It is observed from the dataset that all payments are made by the customers either in the month of due date or later than that. Hence, we consider two parameters—EPID and EPDD as follows:

- (a) EPID = (End of month of invoice due date—Invoice create date)
- (b) EPDD = (End of month of invoice due date—Due date)

## 4 Experimental Analysis

### 4.1 Experimental Setup

In this paper, the forecasting of Account Receivables for a specific month is treated as a two-class classification problem. One class out of these two classes handles those Account Receivables that is credited in a specific month, whereas another handles those Account Receivables that is not credited. Hence, it is considered as a binary classification problem with two class labels—‘yes(1)’ and ‘no(0)’ for the Account Receivables either credited or not credited, respectively, in a specific time span (month).

From the transaction dataset and the general corporate finance [16], we observe that the due invoices are closed within a maximum of 90 day after the due date. Hence, we train three different classification models with an aim to determine whether the invoices will be collected in the month of the due date (current month), in the next month of the due date (next month) or in the next-to-next month of the due date (n2n month). Thus, we calculate the output parameter as—whether the invoices are collected in the current month or not (1 or 0, respectively), the next month or not (1 or 0, respectively), and the next-to-next month or not (1 or 0, respectively) for three models, respectively. Furthermore, for all the three cases—if an invoice is predicted to be closed in the current month of the due date, it is certain that it will be predicted as closed in the next month of the due date too. Similarly, if an invoice is predicted to be closed in the next month of the due date, then it will be predicted as closed in the next-to-next month of the due date too.

For experimental analysis, we perform experiments on aforementioned transaction dataset (Sect. 3.2) for various train data to test data split - 60:40, 70:30, 80:20

**Table 1** Data distribution for classification in current month (80:20 split)

	Collected in current month (class label: 1)	Not collected in current month (class label: 0)	Total
Training data	16,93,964	24,28,880	41,22,844
Testing data	4,23,368	6,07,344	10,30,712
Total	21,17,332	30,36,224	51,53,556

**Table 2** Data distribution for classification in next month (80:20 split)

	Collected in next month (class label: 1)	Not collected in next month (class label: 0)	Total
Training data	22,63,066	18,59,778	41,22,844
Testing data	5,65,392	4,65,320	10,30,712
Total	28,28,458	23,25,098	51,53,556

**Table 3** Data distribution for classification in next-to-next (n2n) month (80:20 split)

	Collected in n2n month (class label: 1)	Not collected in n2n month (class label: 0)	Total
Training data	24,45,042	16,77,802	41,22,844
Testing data	6,10,743	4,19,969	10,30,712
Total	30,55,785	20,97,771	51,53,556

and 90:10. To have a clear insight, we describe the distribution of the data for 80:20 train data to test data split for current month (Table 1), next month (Table 2) and n2n month (Table 3).

Furthermore, the transaction data having a negative amount and/or high DPD is considered as noisy data, and hence, it is highly required to remove this data. For the purpose, we use the Pandas library of Python.

## 4.2 Experimental Results

The key findings from the obtained results (Figs. 2, 3, and 4) are as follows:

- From Figs. 2, 3 and 4, it is easily observed that the logistic regression model gives relatively less accuracy among all the three models. The major limitation of this model is that it keeps linear relation between all the dependent and independent variables. The classification problem addressed in this paper is clearly a nonlinear problem, and hence, the accuracy resulting from logistic regression model is not acceptable.
- From Fig. 3, it is seen that random forest classifier outperforms the logistic regression model. As mentioned (Sect. 3.1.3), random forest classifier fails to remove

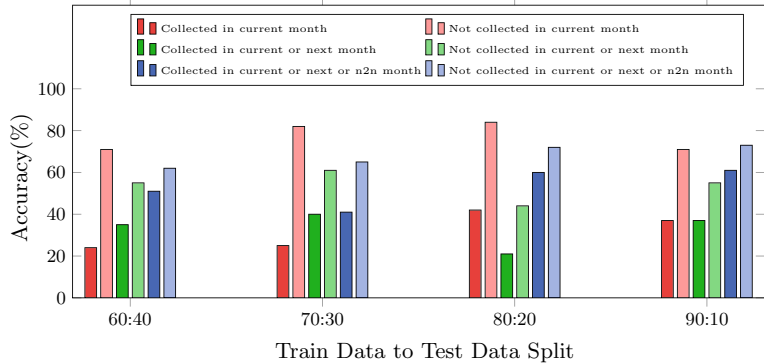


Fig. 2 Accuracy versus train data to test data split (logistic regression model)

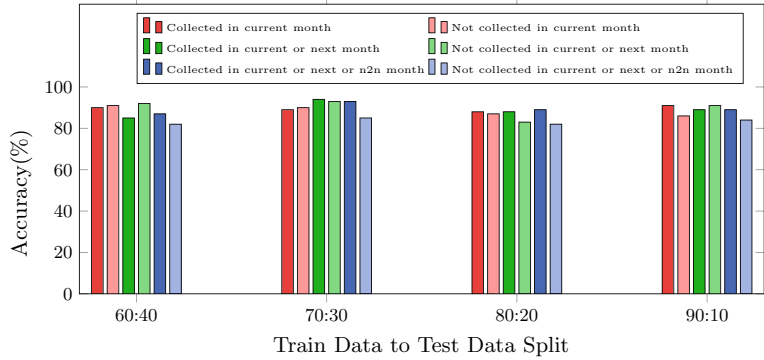


Fig. 3 Accuracy versus train data to test data split (random forest classifier)

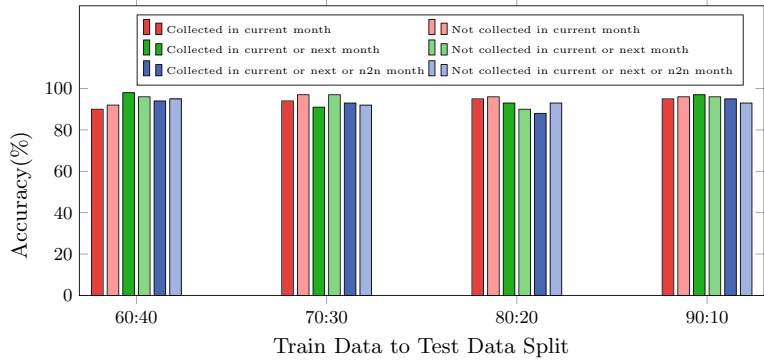


Fig. 4 Accuracy versus train data to test data split (ensemble of K-Means clustering and random forest classifier)

redundancies [12]. Hence, the probability of selecting an irrelevant feature is high because of its random nature of selection (from 100 decision trees taken into account as discussed earlier). This is the reason why a comparatively low accuracy is observed for some categories (Fig. 3).

- From Fig. 4, it is observed that the ensemble of  $K$ -Means clustering and random forest classifier outperforms previously considered two models in all aspects. Here, the key drawback of random forest classifier for not being able to remove redundant data is eliminated by adding  $K$ -Means clustering. The amount and DPD are first passed as parameters in the clustering algorithm as they are the main factors for redundancy that requires to be eliminated. Hence, we observe that the ensemble of random forest classifier and  $K$ -Means clustering results in the highest accuracy among three classification models.

## 5 Conclusion

In this paper, we utilize three machine learning methods to forecast the Account Receivables that any company would receive in a specific time period. We build three models using each of these methods to predict the Account Receivables being closed in (1) the month of the due date, (2) current or next month of the due date and (3) current or next or next-to-next month of the due date. Additionally, we analyse the results obtained through these models for different train data to test data split. From this analysis, we conclude that an ensemble of random forest classifier and  $K$ -Means clustering outperforms among three methods as the data redundancy is handled by  $K$ -Means clustering, and hence, relevant features are utilized for the classifier training. The findings of this paper assist the business executives to forecast the Account Receivables in advance with an aim to gain a better idea of finances and decide strategies for their varied businesses.

## References

1. Abuhommous AA, Mashoka T (2018) A dynamic approach to accounts receivable: the case of Jordanian firms. *Eurasian Bus Rev* 8(2):171–191
2. Maksimovic V (2001) Firms as financial intermediaries: evidence from trade credit data. The World Bank
3. Rajan R, Zingales L (1998) Financial dependence and growth. *Am Econ Rev*
4. Purwanti T (2019) An analysis of cash and receivables turnover effect towards company profitability. *Int J Seocol* 037–044. <https://doi.org/10.29040/seocology.v1i01.6>
5. Appel AP, Malfatti GL, Cunha RLDF, Lima B, de Paula R (2020) Predicting account receivables with machine learning. *arXiv preprint arXiv:2008.07363*
6. Shah H (2019) Customer payment prediction in account receivable. *Int J Sci Res (IJSR)* 8(1):642–644

7. Zeng S, Boier-Martin I, Melville P, Murphy C, Lang CA (2007) Predictive modeling for collections of accounts receivable. In: Proceedings of the 2007 international workshop on Domain driven data mining—DDDM '07, pp 43–48. <https://doi.org/10.1145/1288552.1288558>
8. LaValley MP (2008) Logistic regression. *Circulation* 117(18):2395–2399
9. Pregibon D et al (1981) Logistic regression diagnostics. *Ann Stat* 9(4):705–724
10. Everingham Y, Sexton J, Skocaj D, Inman-Bamber G (2016) Accurate prediction of sugarcane yield using a random forest algorithm. *Agron Sustain Dev* 36(2):27
11. Pappu V, Pardalos PM (2014) High-dimensional data classification. In: Clusters, orders, and trees: methods and applications. Springer, pp 119–150. [https://doi.org/10.1007/978-1-4939-0742-7\\_8](https://doi.org/10.1007/978-1-4939-0742-7_8)
12. Ditterich T (2000) Ensemble methods in machine learning, multiple classifier systems
13. Darbon J, Osher S (2016) Algorithms for overcoming the curse of dimensionality for certain Hamilton-Jacobi equations arising in control theory and elsewhere. *Res Math Sci* 3(1):1–26
14. Aydadenta H, Adiwijaya A (2018) A clustering approach for feature selection in microarray data classification using random forest. *J Inf Process Syst* 14(5):1167–1175
15. <https://www.sailfin.tech/>
16. Bloomenthal A (2020) How long can accounts receivables remain outstanding? <https://www.investopedia.com/ask/answers/021215/how-long-are-accounts-receivable-allowed-be-outstanding.asp>

# Power Supply Board and Switcher Board



**Yash Deep, Madan Chandrashekar, Arivalagan Venkatesan, Shiddalingeshwar Madli, K. A. Manjunath, and Surya Musunuri**

**Abstract** Automotive VLSI industries work on validation of silicon using a hardware board as platform which often requires controlled power supplies and function generator to perform various functionalities. Presently, we are working with bench supplies and bench Function Generators from Various vendors such as Tektronix, Agilent, and Keysight which are very costly and require multiple devices and instruments and multiple host PC's. To make a cost effective and optimized model we have designed boards for power supply to provide power to the hardware board platform and PM Bus for validation of silicon which will cascade multiple power supply boards and hardware board platform for silicon which can be operated using one host PC. By designing the power supply board with inbuilt PMIC and clock, and Switcher board with integrated FTDI USB interface controlled over PM Bus which is designed for 8 channels, the proposed architecture achieves the required various sequence of power and frequency for validating various functionalities of silicon on board. We demonstrate that by using the proposed boards, we need one 220 V, 50 Hz power supply from socket and 1 USB port from PC to control up to 8 validation boards. The board also has current sensors to monitor the current in each power rail and a LCD to display the ratings. We evaluate the proposed architecture by conceptualizing, designing, manufacturing and implementing the boards.

---

Y. Deep (✉) · M. Chandrashekar · A. Venkatesan · S. Madli · K. A. Manjunath · S. Musunuri  
Infineon Technologies India Private Limited, Bengaluru, India  
e-mail: [Yash.Deep@infineon.com](mailto:Yash.Deep@infineon.com)

M. Chandrashekar  
e-mail: [Madan.Chandrashekar@infineon.com](mailto:Madan.Chandrashekar@infineon.com)

A. Venkatesan  
e-mail: [Arivalagan.Venkatesan@infineon.com](mailto:Arivalagan.Venkatesan@infineon.com)

S. Madli  
e-mail: [Shiddalingeshwar.Madli@infineon.com](mailto:Shiddalingeshwar.Madli@infineon.com)

K. A. Manjunath  
e-mail: [Manjunath.Ashwathappa@infineon.com](mailto:Manjunath.Ashwathappa@infineon.com)

S. Musunuri  
e-mail: [Surya.Musunuri@infineon.com](mailto:Surya.Musunuri@infineon.com)

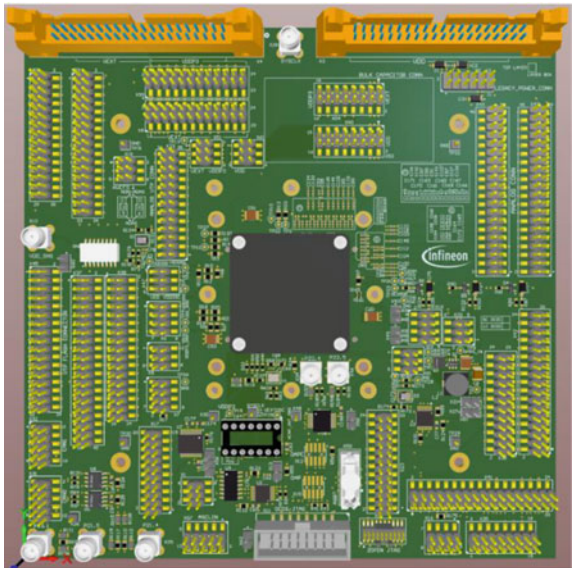


**Keywords** PVTF · Validation boards · PMIC

# 1 Introduction

In majority of the semiconductor industries, the silicon is validated post fabrication for various functionalities and under various conditions before they send out the silicon to the costumers. With Infineon, we have our AURIX 2G automotive microcontroller which we validate. The validation can be divided into two parts- Robustness Validation (RV) and Functional Validation (FV). The FV comprises of validating various IPs inside the silicon. The RV focuses on validating functionality of our automotive microcontroller AURIX™ for all PVTF conditions mentioned in the datasheet. The post silicon validations happens in various validation platforms like PCB boards and software tools. One such hardware validation platforms is ROVACS board (Fig. 1) built within Infineon Technologies. The ROVACS Board is used to validate multiple functionalities of a silicon ranging from DIO, PORTS, CAN, GETH, LVDS characterization, speed monitoring etc. and is mainly used for RV. To provide clean and noise free input power, we have two 40 pin connector on the board to provide voltages. These voltages are provided via bench supplies from Tektronix, Agilent, Keysight. The clock is fed via the function generators to the SMA connectors of the board.

**Fig. 1** ROVACS board with two 40 pin power connector (top) and SMA clock connector



1.1 PVTF Conditions

The PVTF stands for Process, Voltage, Temperature and Frequency. Different splits of Silicon are fabricated. Every split contains different p-type and n-type doping. Less doping means less charge carriers i.e. slow. More doping means more charge carriers i.e. fast. With that concept, splits are fabricated as SS, SF, POR, FS, FF (process corners). For example: SF split implies slow for p-type and fast for n-type. We use these splits to validate silicon functionality with deviations from the optimum doping that can be caused during fabrication.

Aurix™ 2G requires 3 different voltage levels, 5, 3.3 and 1.25 V. These voltages are supplied to the ROVACS board via the bench supplies. Each rail is varied from ± 20% of the nominal voltage value. We see if the test flashed on Si is failing/passing for each combination of the three.

Temperature of the Silicon is changed with the help of thermo-stream which is placed over the board. The temperature is varies from 40 to 170 °C (the automotive standards).

Nominal frequency for A2G Aurix™ is 300 MHz. We connect the frequency generator output to Board’s crystal pin and changes the frequency on the frequency generator. This constitutes our PVTF validation conditions. The results are then plotted in the schmoo chart (Fig. 2). The X-axis and Y-axis can represents voltage, frequency or temperature.

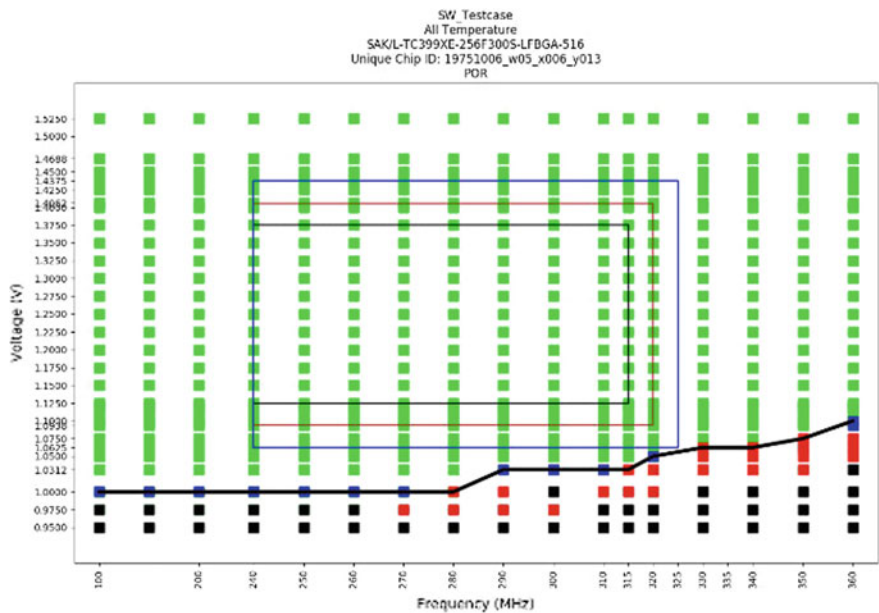


Fig. 2 Schmoo chart. Green-PASS, red-fail, black-error access or connection issue, blue-beyond datasheet specification

## 1.2 *PM Bus*

The Power Management Bus (PM Bus) [1], is a variant of the System Management Bus (SM Bus) which is targeted at digital management of power supplies. Like SM Bus, it is a relatively slow speed two wire communications protocol based on IIC. Unlike either of those standards, it defines a substantial number of domain-specific commands rather than just saying how to communicate using commands defined by the reader. 400 kHz speeds are allowed in PM Bus. It only supports 7 bit addressing.

For the faster validation process, multiple boards are used together. Each board requires 2 bench supplies and 1 function generator. As we head towards automation of these processes, this would lead to blown out procurement costs for the bench supplies and function generator.

In this paper we propose two different boards which can be used to automate and scale the validation process and are cheaper alternatives to the bench power supplies and function generator. These boards work on the PM Bus topology to send and receive the data.

## 2 **Implementation**

To overcome the issues we came up with 2 boards which can lead to automation of the validation process and optimize the cost.

### 2.1 *Power Supply Board*

The Power supply board is designed by keeping the following points in mind-

- Should support current rating demands of Aurix™ 2G Microcontroller.

- Voltage rails/ clock source necessary for Aurix™ 2G Microcontroller.

- Power rails/ clock sources to be variable and controlled by a host PC.

- On board current sensing and display options.

- Should also work as a standalone board without any programming.

The Power Supply Board offers a low cost solution for powering any validation boards (ROVACS in our case).

The Power Supply Board takes in 12 V as input from DC jack and generates 1.25, 3.3 and 5 V using the PMIC TPS65400 [2]. It has 4 channels, of which 3 are used to power up the ROVACS board. All three power rails can be varied by changing the Vref value of the PMIC via the PM Bus. The 1.25 V can be varied from 0.8 to 1.87 V, 3.3 V can be varied from 1.38 to 4.3 and 5 V can be varied from 2.1 to 6.54 V. This is application specific and can be changed from the resistor network implemented on the board. The 4th additional channel works at 5 V with default Vref. 1.25 and 5 V

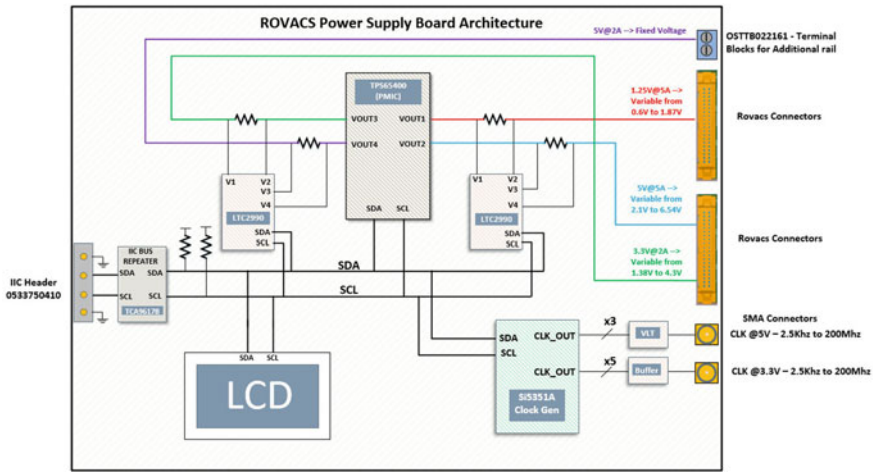


Fig. 3 Power supply board architecture

channel has a current rating of 5 A and 3.3 V rail has a current rating of 2 A which is sufficient for our AURIX microcontroller.

The board also has 2 current sensors which supports current sensing in mA resolution, 2 channels each, to calculate the current consumption in each rail. The resistor values are calculated according to the max voltage and current swing and are placed in series as shown in Fig. 3. The data is read via the PM Bus. A LCD is also added to the PM Bus to show the current ratings in real time.

The Power Supply Board has a clock source implemented with Si5351A. It has 8 channels going out to the SMA connector for clock output to be fed to the microcontroller. The frequencies can be varied from 2.5 kHz to 200 MHz via the PM Bus. 3 of the clock channels work at 5 V and other 5 channels are level shifted to 3.3 V.

The PM Bus is implemented on the board with different slave devices connected to it. The end of PM Bus goes to an IIC bus repeater and then to a 4 pin header sharing the PM Bus with Switcher board. Appropriate pull up resistor values are calculated using formula

$$R_p(\min) = \frac{V_{cc} - V_{ol}(\max)}{I_{ol}} \quad (1)$$

where  $R_p$  [3], is the pull up resistor value,  $V_{cc}$  is the pull up reference voltage,  $V_{ol}$  is the low level output voltage and  $I_{ol}$  is low level output current. The maximum pull-up resistance is limited by the bus capacitance ( $C_b$ ) due to I2C standard rise time specifications. If the pull-up resistor value is too high, the I2C line may not rise to a logical high before it is pulled low. The maximum value of  $R_p$  is calculated using the formula below

**Fig. 4** Power supply board

$$R_p(\max) = \frac{T_r}{0.8473 \times C_b} \quad (2)$$

where  $T_r$  is the rise time and  $C_b$  is the bus capacitance.

The board also has 2 DC jack, one as input and the other as output. This is done to cascade multiple Power Supply boards together and power multiple boards with only one 220 V DC jack (Fig. 4).

The board is a 6 layer board, with total number of components exceeding 300 and with 695 connections.

## 2.2 Switcher Board

The switcher board establishes a connection between the host PC and the power supply board via the PM Bus.

The Switcher Board (Fig. 5) has a USB jack which is to be connected to the host PC. The board also has FT2232HL [4]. The FT2232HL is a USB 2.0 High Speed (480 Mb/s) to UART/FIFO IC. It has the capability of being configured in a variety of industry standard serial or parallel interfaces. It has dual Multi-Protocol Synchronous Serial Engine (MPSSE) to simplify synchronous serial protocol (USB to JTAG, I2C, SPI or bit-bang) design.

The Board also has TCA9548A [5], IIC switch. The TCA9548A device has eight bidirectional translating switches that can be controlled through the IIC bus. The

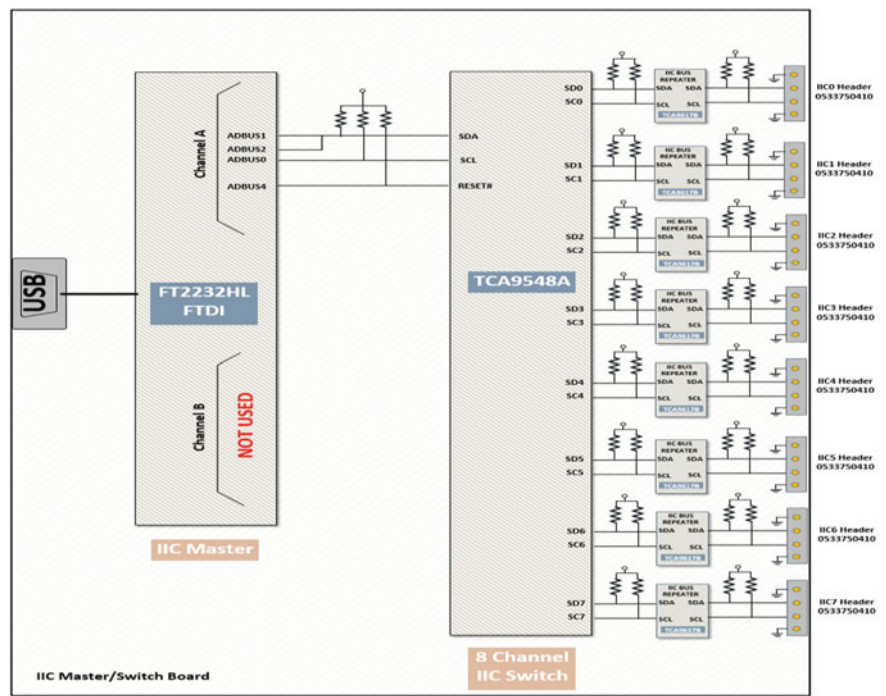


Fig. 5 Switcher board architecture

SCL/SDA upstream pair fans out to eight downstream pairs, or channels. Any individual SCn/SDn channel or combination of channels can be selected, determined by the contents of the programmable control register. This acts as a IIC switch as shown in Fig. 4.

Each of the 8 channels are passed through TCA9617B [6], IIC Bus Repeater. The TCA9617B is a BiCMOS dual bidirectional buffer intended for IIC bus and PMBus systems. It can provide bidirectional voltage-level translation (up-translation and down-translation) between low voltages (down to 0.8 V) and higher voltages (2.2–5.5 V) in mixed-mode applications. This device enables IIC and similar bus systems to be extended, without degradation of performance. In the end each of these channels are ended with a 4 pin connector compatible with Power Supply Board.

The Switcher board is used to control all the slave devices present within the PM Bus. This includes changing the voltage of PMIC and changing the frequency of the clock generator in Power Supply Board (Fig. 6).

The Switcher board is a 4 layer board with 150 components and 430 connections.

**Integration.** To use the Power Supply Board and Switcher board with the ROVACS board, we need a 12 V adapter for powering up the boards, two 40 wire ribbon cable for connecting Power Supply board to the ROVACS board and a USB cable for connecting Switcher board to the host PC (Fig. 7).



Fig. 6 Switcher board

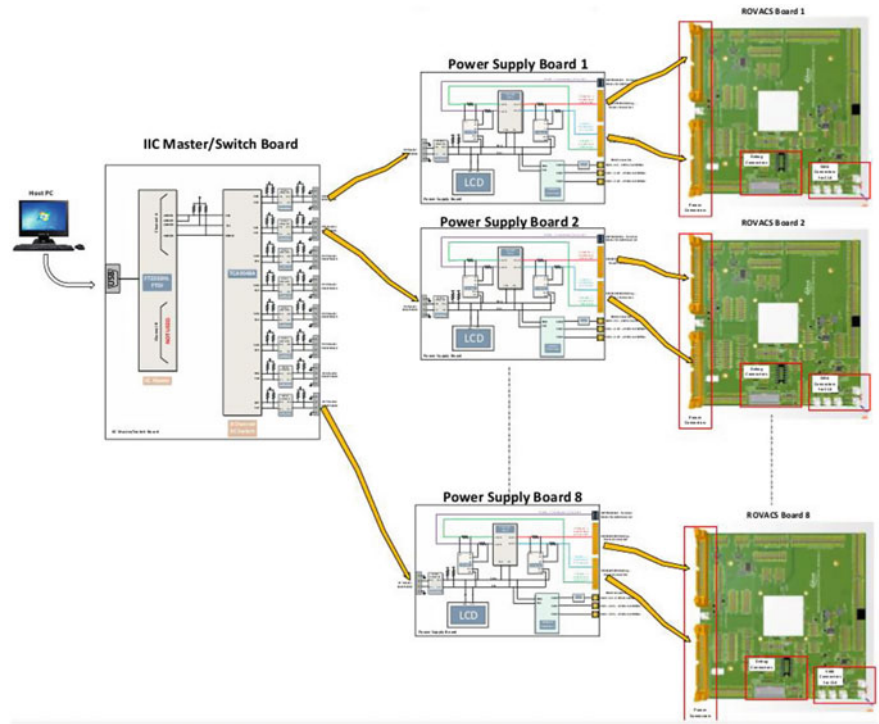
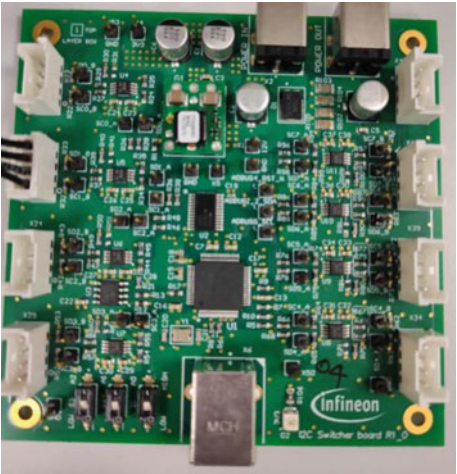


Fig. 7 Integrating power supply board and switcher board

One port of the host PC controls 1 Switcher board which has 8 channels. Each channel is connected to 1 Power Supply Board. Each Power Supply Board is connected to 1 ROVACS validation board.

To vary the voltage or frequency of a particular ROVACS board, the host PC first sends the data to the IIC switcher to select the channel with which it wants to communicate. After selecting the channel, it sends out the slave address of the PMIC or function generator (or any other save device). Since all the slave devices is connected in the PM Bus, each has its own slave address. After receiving the acknowledgment, the host PC sends out the change to the particular slave, resulting in the changing parameters of the intended ROVACS validation platform.

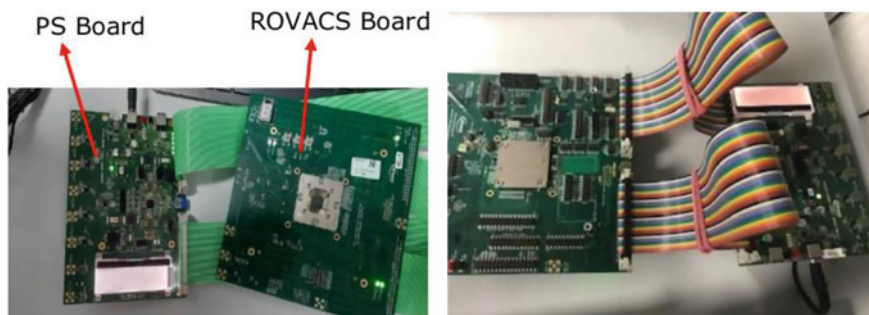
### 3 Discussion

The Power Supply board along with the Switcher boards have been conceptualized, designed, manufactured and implemented (Fig. 8).

The Power supply board can also function as a standalone board to power up the ROVACS validation platform hence eliminating 2 bench supplies as shown in Fig. 9.

#### 3.1 Cost Considerations

For the automation of validation process, we required 30 ROVACS boards set up in our labs with 30 Power supply board and 8 Switcher boards. In contrast, when we do not use these setups, we will require 60 bench supplies for 30 ROVACS validation boards. The total cost for developing these boards have been given in Table 1. Table 2 shows the procurement cost of 60 bench supplies.



**Fig. 8** PS board connected to ROVACS (left), 40 wire ribbon cable for connecting the ROVACS (right)





**Fig. 9** Multiple ROVACS powered using PS boards

**Table 1** Development cost of PS board and switcher board

	Power supply board	Switcher board
NRE cost	776€	591.7€
No. of boards	30	8
Cost/board	383.15€	287.12€
Total board cost	11,494.5€	2296.96€
Effort cost	4000.00€	4000.00€
Total cost	16,270.5€	6888.66€

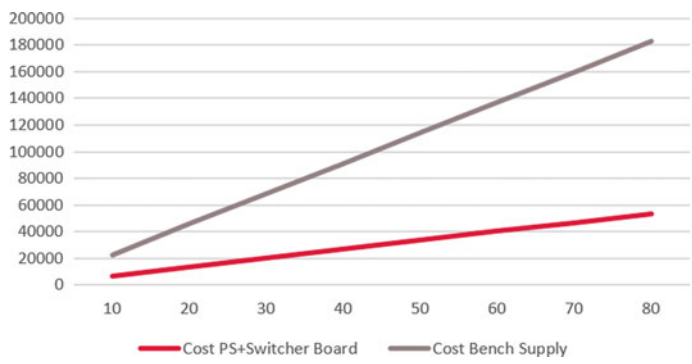
**Table 2** Procurement cost of bench supplies

Model	Cost/unit
E3633A cost/power supply	1141.82€
No. of power supplies	60

The total cost of developing Power supply board and switcher board comes out to be 23,159.16€.

The total cost of procuring 60 bench supplies comes out to be 68,509.03€. The total cost saving hence comes out to be around 45,349.87€.

Even if we neglect all the one-time cost and only consider the cost of the boards, we have a saving of 1613.37€ per ROVACS validation setup. If we use the Power Supply board as a stand-alone to replace the bench supplies, we have a cost saving of 1900.49€.



**Fig. 10** Graph showing the cost differences between the PS + Switcher board and bench supplies

Figure 10 shows the cost difference between the power supply plus switcher board and bench supplies when we scale the setup. The X-axis shows the total number of boards and Y-axis shows the amount in EUR.

4 Results

All the functionalities of the Power supply board and Switcher Board have been verified and tested.

Table 3 shows the expected and measured voltage from the PMIC. The resolution of VDD, VEXT and VDDP3 is measured 10, 35 and 25 mV approximately. Apart from TPS65000 PMIC, we are able to access the I2C slaves, I-Sense LTC2990 and LCD NHD-C0220BiZ. These functionalities altogether will mimic a typical triple output power supply.

Figure 11 shows the output of 2 channels, configured at 21 MHz and 24 MHz. The clocks show jitteriness behavior that is possibly to get eliminated by non-fractional div/spread spectrum enable and other options from Si5351b.

With this we will be able to sweep voltage and speed input of ROVACS or any other validation platform. Multiple clock settings from 8 clock out pins with different multiplier/divider can be fed to ROVACS and speed performance is used to be correlate with available RV speed shmoo result.

**Table 3** Output voltage of the PMIC

PMIC_CH	Expected vout (V)	Measured board (V)
CH1_VDD	1.25	1.25
CH2_VEXT	5.04	4.99
CH3_VDDP3	3.312	3.27

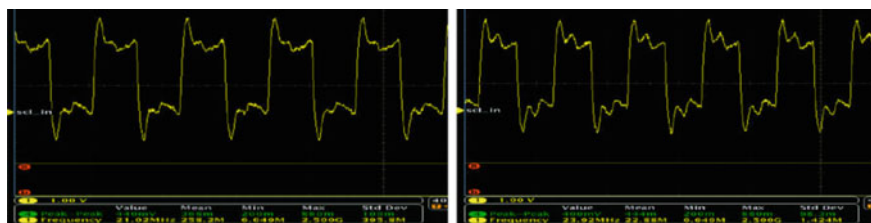


Fig. 11 Clock output 21 MHz (left) and 24 MHz (right)

## 5 Conclusion and Future Work

We presented a methodology which is cost efficient, scalable and automated which is implemented in hardware validation platforms. The methodology can also be applied in future generations of silicon. We started with a discussion on hardware validation platforms and what are the power and clock requirements for it. Then we looked into the importance of varying voltage and frequency for PVT conditions. We presented 2 boards architectures which can be used to replace the costly function generators and bench supplies. We applied our architecture in PCB and tested out different modules. We also tested out the feasibility by connecting the whole setup with ROVACS board and carried out various validation codes. We concluded with how 8 different validation boards can be controlled via 1 host PC USB port and one 220 V power source. We also showed the cost reduction with our setup in the end.

The methodology demonstrated can be extended to various hardware validation platforms. Then we intend to apply this methodology to our next generation of microcontroller. We are also working on automating the whole setup by having the python wrappers ready to change the voltage and frequency automatically.

## References

1. [https://en.wikipedia.org/wiki/Power\\_Management\\_Bus](https://en.wikipedia.org/wiki/Power_Management_Bus)
2. [https://www.ti.com/lit/ds/symlink/tps65400.pdf?ts=1609948426568&ref\\_url=https://www.google.com/](https://www.ti.com/lit/ds/symlink/tps65400.pdf?ts=1609948426568&ref_url=https://www.google.com/)
3. [https://www.ti.com/lit/an/slua689/slua689.pdf?ts=1613571684857&ref\\_url=https://www.google.com/](https://www.ti.com/lit/an/slua689/slua689.pdf?ts=1613571684857&ref_url=https://www.google.com/)
4. [https://www.mouser.in/datasheet/2/163/DS\\_FT2232H-1621240.pdf](https://www.mouser.in/datasheet/2/163/DS_FT2232H-1621240.pdf)
5. [https://www.ti.com/lit/ds/symlink/tca9548a.pdf?ts=1613551479940&ref\\_url=https://www.google.com/](https://www.ti.com/lit/ds/symlink/tca9548a.pdf?ts=1613551479940&ref_url=https://www.google.com/)
6. [https://www.ti.com/lit/ds/symlink/tca9617b.pdf?ts=1613553892150&ref\\_url=https://www.google.com/](https://www.ti.com/lit/ds/symlink/tca9617b.pdf?ts=1613553892150&ref_url=https://www.google.com/)

# Canonical Polyadic Decomposition of EEG Image Tensor for BCI Applications



K. Keerthi Krishnan and K. P. Soman

**Abstract** EEG signals represented as multidimensional array with dimensions of time, frequency, and tasks are being explored for the last few decades for BCI applications. In this work, a method to differentiate distinct components of these dimensions from EEG signals for the four motor imaginary tasks is explored. The EEG signal spectrum computed using STFT of the VMD modes of each trial is arranged in the form of a tensor. Further, CP decomposition is employed on EEG image tensor with modes as time, frequency, and tasks. The results illustrate that the method provides distinguishing components for each mode. Hence, the method can be adopted for analyzing the EEG signals for identifying and differentiating the components when expressed as a multidimensional array.

**Keywords** EEG · VMD · Tensor decomposition · CP-ALS · BCI · EEG image

## 1 Introduction

Electroencephalography (EEG) is a contemporary way to examine the brain conditions as well as a tool to develop mechanisms for thought recognition. In the case of Brain Computer Interface (BCI), the EEG signal features in the space–time, frequency–time, and space–frequency–time have been explored since few decades [9]. Since EEG signals are inherently multidimensional which exhibits variations across channels, trials as well as among the subjects, an optimal combination of time–frequency and spatial analysis is required. Tensor decomposition is a method involving decomposition of multidimensional signals and has manifested as an effective way in applications like data analysis, blind source separation, in addition to EEG signal processing [2]. Tensor decomposition finds wide applications in analysis of hyperspectral images [16]. The two popular methods for tensor decomposition are canonical polyadic (CP) or parallel factor analysis (PARAFAC) [6] and Tucker [18].

---

K. K. Krishnan (✉) · K. P. Soman

Center for Computational Engineering Networking (CEN), Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

e-mail: [kp\\_soman@amrita.edu](mailto:kp_soman@amrita.edu)

The tensor decomposition of EEG signals are applied for source localization, cognitive neuroscience, clinical neuroscience, and BCI applications [3, 20]. Meaningful sources are obtained from CP decomposition of coupled magnetoencephalography (MEG) and EEG tensors represented using Morlet wavelet decomposition [10, 11]. EEG signals spectrum is converted into image form with STFT of empirical modes after decomposing using Empirical Mode Decomposition (EMD) [13]. Variational mode decomposition (VMD) which iteratively decomposes real-valued multimodal signal like EEG by applying calculus of variation [4, 17]. The VMD method of finding modes with frequency support has been applied on EEG signals [12, 15].

The objective of this study is to identify distinct components from simultaneous multiclass time-varying EEG signal spectrum in the form of image. The EEG signals on decomposition provide four VMD modes. The STFT of each of these modes is computed and concatenated to form image. The EEG image tensor is formed by arranging the spectrum of EEG signals of the four different motor imaginary tasks as the third dimension. CPD is employed on the EEG image tensor. The components obtained after CP decomposition are analyzed using correlation. This paper is organized as follows: first the details of dataset and methods are described. Then, results and discussion in which the development of tensor components and the relevance are discussed.

## 2 Dataset

BCI competition Dataset IIIa [1] consists of EEG recordings of three subjects captured using 60 channels. The EEG signal readings are captured from 60 trials per each subject. In each trial of a subject, the EEG signals are taken when the subject imagined the movement of either of the following four imaginary movements. The movements being the imagination of moving left hand, right hand, foot, or tongue. Each trial is of duration 10 s, during which imagination was executed for 5 s starting from 3 s. The imagination was prompted when an arrow appeared on the screen in any one of the four directions, after which the subject imagines the corresponding movement till the arrow disappeared. The EEG signals were sampled at 250 Hz and filtered between 1 Hz and 50 Hz.

## 3 Methodology

### 3.1 Variational Mode Decomposition

VMD breaks down the real multimodal signal adaptively, into number of modes. The computation involves optimization methodology which is solved iteratively using alternate direction method of multipliers (ADMM). The modes obtained are centered

around frequency,  $\omega_l$ . The bandwidth of each mode is minimized subject to the condition that the original signal is equal to the sum of the obtained modes [4]. For a real signal  $s(t)$ , the process can be mathematically expressed as

$$\min_{s_l, \omega_l} \left\{ \sum_l \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * s_l(t) \right] e^{-j\omega_l t} \right\|_2^2 \right\} \text{s.t. } \sum_l s_l = s \quad (1)$$

where  $s_l(t)$  is the  $l$ (th) mode with center frequency  $\omega_l$ . Also,  $\delta$  is the Dirac delta and  $*$  denotes the convolution. The augmented Lagrangian equation is formulated along with a quadratic penalty term, balancing parameter of data fidelity constraint,  $\alpha$  and Lagrangian multiplier,  $\lambda$ . The Lagrangian equation is written as

$$\begin{aligned} \mathcal{L}(s_l, w_l, \lambda) = & \alpha \sum_l \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * s_l(t) \right] e^{-j\omega_l t} \right\|_2^2 \\ & + \left\| s(t) - \sum_l s_l(t) \right\|_2^2 + \left\langle \lambda(t), s(t) - \sum_l s_l(t) \right\rangle \end{aligned} \quad (2)$$

Equation 2 has a solution computed with ADMM as shown in Eqs. 3 and 4.

$$\hat{s}^{n+1}(\omega) \leftarrow \frac{\hat{x}(\omega) - \sum_{i < l} \hat{s}_i^{n+1} - \sum_{i > l} \hat{s}_i^n + \frac{\lambda^n}{2}}{1 + 2\alpha(\omega - \omega_l^n)^2} \quad (3)$$

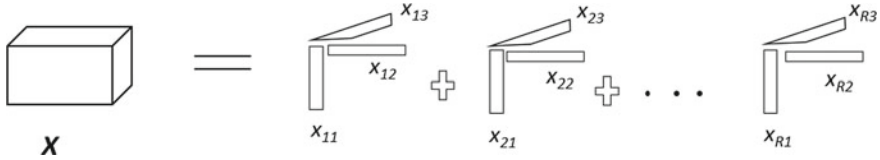
$$\omega_l^{n+1} \leftarrow \frac{\int_0^\infty \omega \left| \hat{s}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{s}_l^{n+1}(\omega) \right|^2 d\omega} \quad (4)$$

where the capitalized letters represent Fourier transform of  $s(t)$ .

### 3.2 Tensor Decomposition

CP decomposition on tensor is a higher-order singular value decomposition that breaks down tensor into multiple components. The components after CPD will be rank one tensors whose sum gives back the tensor. In the case of three-way tensor,

$X \in R^{K_1 \times K_2 \times K_3}$ , the decomposition provides  $X \approx \sum_{r=1}^R x_{r1} \circ x_{r2} \circ x_{r3}$  where  $R$  is the rank of the tensor,  $\circ$  represents outer product,  $x_{1r} \in R^{K_1}$ ,  $x_{2r} \in R^{K_2}$ ,  $x_{3r} \in R^{K_3}$  for rank,  $r = 1$  to  $R$ .  $x_{r1}$ ,  $x_{r2}$ ,  $x_{r3}$ , represents  $r$ th column of the component matrix,  $X_1 \in R^{K_1 \times R}$ ,  $X_2 \in R^{K_2 \times R}$ ,  $X_3 \in R^{K_3 \times R}$ , respectively. Figure 1 illustrates the tensor and the component matrices using CPD.



**Fig. 1** CP tensor decomposition into Rank 1 components

Mode-1 fibers are linear combination of columns of  $X_1$ . Similarly, mode-2 and mode-3 fibers are linear combination of columns of  $X_2$  and  $X_3$ , respectively. The component matrices are computed as optimisation problem,

$$\min_X \|X - \hat{X}\| \text{ where } \hat{X} = \sum_{r=1}^R x_{r1} \circ x_{r2} \circ x_{r3} \quad (5)$$

Among the various optimization algorithms for CP decomposition, such as enhanced line search (ELS), alternating least squares (ALS), alternating slice-wise diagonalisation (ASD), and self-weighted alternating trilinear decomposition (SWATLD), ALS remains as one of the superior method [5, 7]. The ALS approach involves computation of Khatri–Rao product of  $X_2$  and  $X_3$  to solve for  $X_1$ , then Khatri–Rao product of  $X_1$  and  $X_3$  to solve for  $X_2$ , then Khatri–Rao product of  $X_1$  and  $X_2$  to solve for  $X_3$ , and procedure is repeated until convergence criterion is satisfied. Khatri–Rao product of two matrices reduces the problem to a linear least squares problem as given below

$$\hat{X}_1 = \min_{X_1} \|X_{(1)} - X_1(X_2 \odot X_3)^T\|_F \quad (6)$$

The optimal solution is given by

$$\hat{X}_1 = X_{(1)}(X_2 \odot X_3)(X_2^T X_2 \odot X_3^T X_3)^\dagger \quad (7)$$

where  $\dagger$  represents Moore–Penrose pseudoinverse. Similarly,  $\hat{X}_2$  and  $\hat{X}_3$  are computed. The final solution is obtained by iterating the procedure till convergence criterion is satisfied.

## 4 Results and Discussions

EEG signals for a duration 5 s during which the subject imagined the movement were decomposed into 4 VMD modes. In order to identify the different frequency bands constituting the EEG signals, EEG signal from each electrode was separately broken down into 4 VMD modes. The power distribution of each mode was observed using

the STFT. For computing the STFT, a window length of 200 milliseconds with an overlap of 75% is provided. The Gabor window was used as the window function. The STFT which provides spectrogram of all 4 modes is concatenated to form an EEG image for an electrode [8]. The EEG image for subject 1 and for a single electrode is shown in Fig. 2.

Further in order to analyze the embedded information simultaneously in the different channels and tasks, EEG images were represented as tensor. The EEG images were concatenated as 3D tensor. The details of different tensor formulation are provided in Fig. 3a. The tensor was formed by stacking the EEG image of four different classes, viz. imaginary movement of right hand, left hand, foot, and tongue. The dimension of the tensor is  $124 \times 81 \times 4$ .

The rank of the tensor is computed based on the relative error. The smallest integer where relative error is smaller than the maximum real error is taken as the rank of the tensor while lower bound is computed using the truncation error of the tensor's multilinear singular values [19]. The variation of lower bound for different ranks is shown in Fig. 3b, where a rank of 9 is chosen.

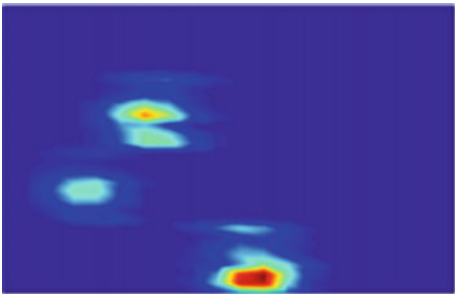


Fig. 2 EEG image formed using STFT of VMD modes

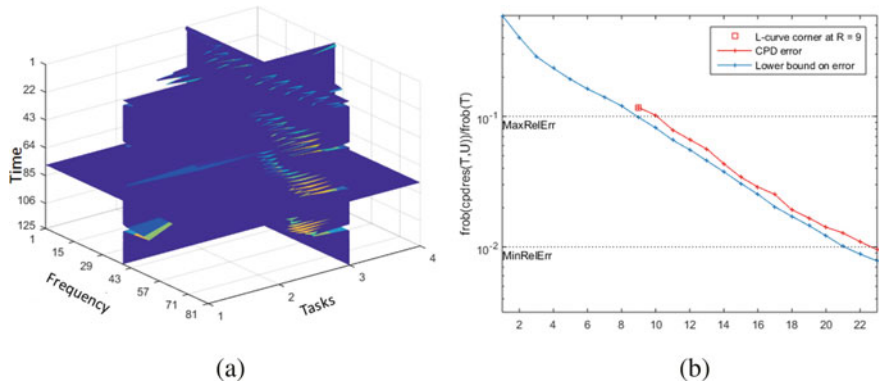
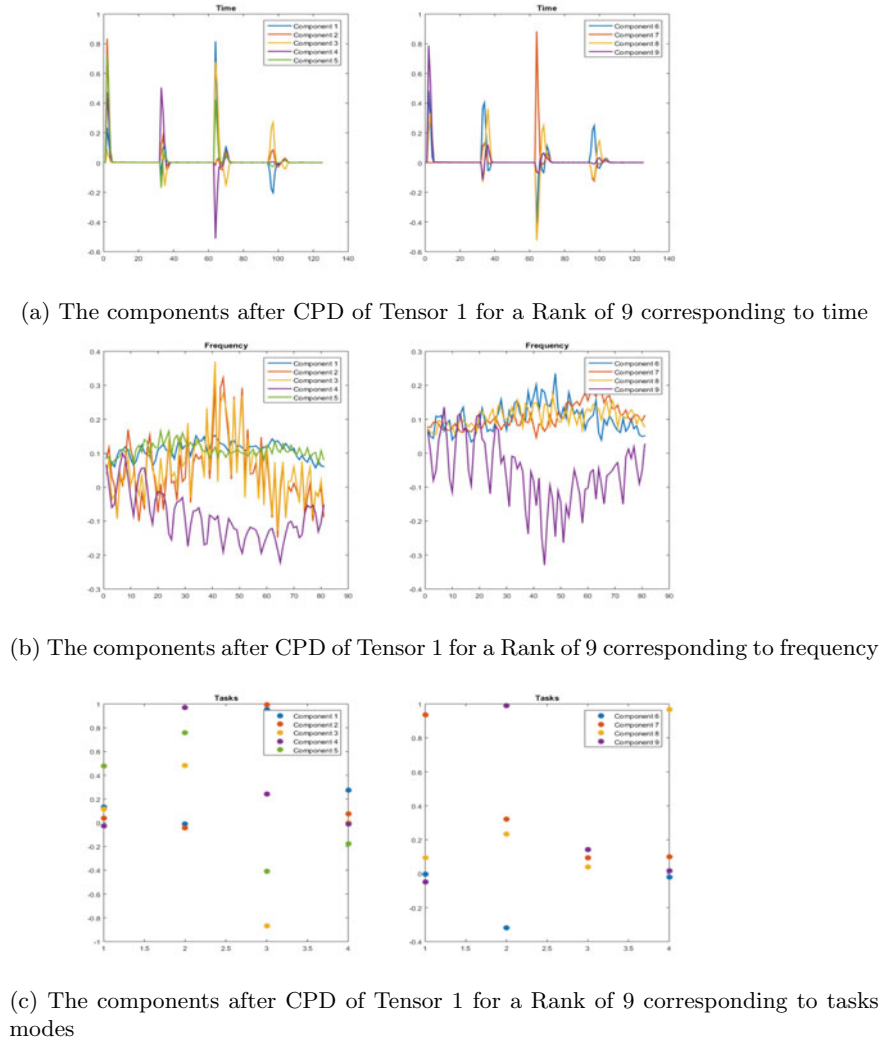


Fig. 3 a Tensor formation using EEG image b rank computation





**Fig. 4** a–c shows the components after CPD of Tensor 1 for a Rank of 9

The CPD of EEG image tensor is computed with ALS algorithm. The components were computed iteratively till the fitness function is 9.1. The columns of factor matrices represent the signatures of the underlying components for the corresponding dimension. These 9 columns or components for each of the 3 modes of the tensor are plotted in Fig. 4. Figure 4a–c shows the time, frequency, and task modes for subject 1 for a single electrode. It can be observed that each component peak varies both in the time and frequency modes. From Fig. 4a, it can be observed that the component

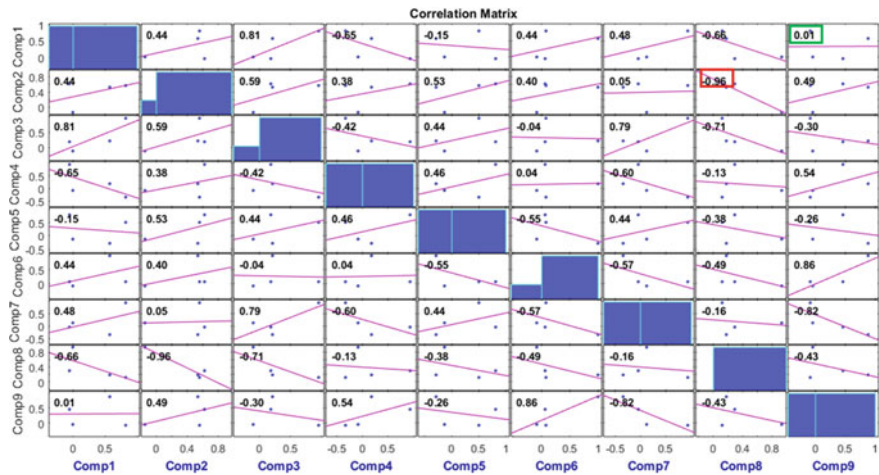


Fig. 5 Pearson correlation coefficient for the nine components of task

3 of time mode has a peak in 65 while from Fig. 4b highest frequency component at 40 for component 3.

In order to evaluate how much variation is present in each components compared to others, Pearson correlation coefficient of the different components is computed for each of the modes. The Pearson correlation coefficient for components in the task mode is plotted in Fig. 5. The highest correlation is marked in red color, and the lowest is marked in green color. Similarly, the highest correlation of 0.98 is observed in the case of time mode for the components 2 and 4, while in the case of frequency mode all correlation coefficient remains below 0.78. The discriminating components are obtained from PARAFAC decomposition of EEG signals with time–frequency representations [14] with wavelets and EMD modes.

## 5 Conclusions

A novel method of tensor formation using EEG image is proposed. On applying CPD on the tensor having time, frequency, and tasks as the dimensions, it is observed that the distinctive components are obtained in each mode. The discussion shows that the method can be applied for identifying the tasks in the case of BCI applications. Further, this method can also be adopted for EEG source localisation where simultaneous monitoring of the multiple factors is required. Even though the method provides unique techniques for multi-way analysis of EEG signals, the computational requirements are higher when compared to spectrum computation using wavelet transforms and STFT of EEG signals.

## References

- Blankertz B, Muller KR, Krusienski DJ, Schalk G, Wolpaw JR, Schlogl A, Pfurtscheller G, Millan JR, Schroder M, Birbaumer N (2006) The bci competition iii: Validating alternative approaches to actual BCI problems. *IEEE Trans Neural Syst Rehab Eng* 14(2):153–159
- Cichocki A, Mandic D, De Lathauwer L, Zhou G, Zhao Q, Caiafa C, Phan HA (2015) Tensor decompositions for signal processing applications: from two-way to multiway component analysis. *IEEE Signal Process Mag* 32(2):145–163
- Cong F, Lin QH, Kuang LD, Gong XF, Astikainen P, Ristaniemi T (2015) Tensor decomposition of eeg signals: a brief review. *J Neurosci Methods* 248:59–69
- Dragomiretskiy K, Zosso D (2013) Variational mode decomposition. *IEEE Trans Signal Process* 62(3):531–544
- Faber NKM, Bro R, Hopke PK (2003) Recent developments in candecomp/parafac algorithms: a critical review. *Chemomet Intell Lab Syst* 65(1):119–137
- Hitchcock FL (1927) The expression of a tensor or a polyadic as a sum of products. *J Math Phys* 6(1–4):164–189
- Kolda TG, Bader BW (2009) Tensor decompositions and applications. *SIAM Rev* 51(3):455–500
- Krishnan KK, Soman K (2021) CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image. *Biomed Eng Lett* 1–13
- Miwakeichi F, Martinez-Montes E, Valdés-Sosa PA, Nishiyama N, Mizuhara H, Yamaguchi Y (2004) Decomposing eeg data into space-time-frequency components using parallel factor analysis. *NeuroImage* 22(3):1035–1045
- Mørup M, Hansen LK, Herrmann CS, Parnas J, Arnfred SM (2006) Parallel factor analysis as an exploratory tool for wavelet transformed event-related eeg. *NeuroImage* 29(3):938–947
- Naskovska K, Lau S, Korobkov AA, Haueisen J, Haardt M (2020) Coupled CP decomposition of simultaneous MEG-EEG signals for differentiating oscillators during photic driving. *Front Neurosci* 14
- Pandey P, Seeja K (2019) Subject independent emotion recognition from EEG using vmd and deep learning. *J King Saud Univ Comput Inf Sci*
- Park C, Looney D, ur Rehman N, Ahrabian A, Mandic DP (2012) Classification of motor imagery bci using multivariate empirical mode decomposition. *IEEE Trans Neural Syst Rehab Eng* 21(1):10–22
- Pouryazdian S (2017) Tensor analysis of electroencephalogram signal for localization of event-related potentials. Ph.D. thesis, Ryerson University
- Rahman MM, Bhuiyan MIH, Das AB (2019) Classification of focal and non-focal EEG signals in YMD-DWT domain using ensemble stacking. *Biomed Signal Process Control* 50:72–82
- Renu R, Sowmya V, Soman K (2018) Pre-processed hyperspectral image analysis using tensor decomposition techniques. In: *International symposium on signal processing and intelligent recognition systems*, pp 205–216. Springer
- Soman K, Kumar SS, Mohan N, Poornachandran P (2019) Modern methods for signal analysis and its applications. In: *Recent advances in computational intelligence*, pp 263–290. Springer
- Tucker LR (1966) Some mathematical notes on three-mode factor analysis. *Psychometrika* 31(3):279–311
- Vervliet N, Debals O, Sorber L, Van Barel M, De Lathauwer L (2016) Tensorlab user guide. Available on: <http://www.tensorlab.net>
- Williams AH, Kim TH, Wang F, Vyas S, Ryu SI, Shenoy KV, Schnitzer M, Kolda TG, Ganguli S (2018) Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron* 98(6):1099–1115

# Generalized Constraint Representation of Adjectives in PNL



Bushra Siddique and M. M. Sufyan Beg

**Abstract** Natural languages (NLs) are complex in nature. Easily understood by humans, it is challenging to make machines understand and process NLs. While the traditional paradigm of computing with numbers falls short in dealing with NL-based information, Zadeh's paradigm of Computational Theory of Perceptions (CTP) serves to deal with it. Based on this theory, the concept of Precisiated Natural Language (PNL) proposes to precisiate the meaning of NL propositions. To be able to carry out meaning precisiation in PNL, one of the core issues is constraint explicitation, i.e., to be able to explicitate the implicit constraints of these NL propositions. This paper addresses this subject in the scope of the lexical category of 'adjectives'. Since the noun phrases consisting of nouns and adjectives constitute a significant amount of information in NL-based data, to be able to explicitate the constraints of such noun phrases would result into a fair amount of precisiated information on which computation could be carried out. In this paper, we present the semantics of adjectives in light of fuzzy variables, propose a generalized constraint representation suitable for adjectives and report the implementation of the same.

**Keywords** Computing With Words · CWW · Adjectives · PNL · Precisiated Natural Language · Generalized constraint

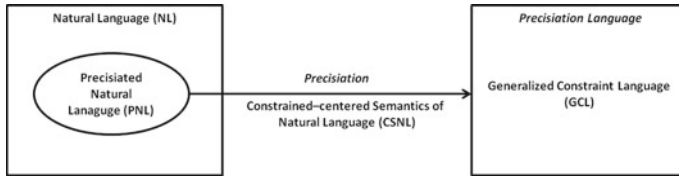
## 1 Introduction

The paradigm of Computational Theory of Perceptions (CTP) is based on Zadeh's paradigm of Computing With Words (CWW) [10]. In this paradigm, the objects of computation are words rather than numbers. Thus, it is able to address the limitations of the traditional paradigm of computing in dealing with natural language and makes this imprecise information accessible to the machines. This CWW concept is based on the semantics of Natural Language (NL) and Fuzzy Logic Theory (FLT).

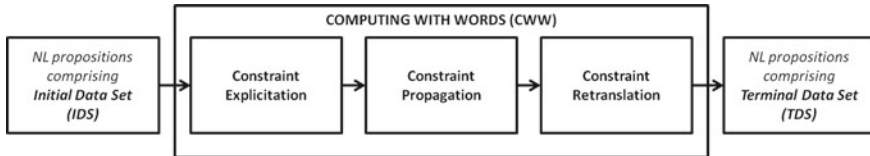
The underlying idea of CTP is to translate the natural language statements/phrases in a form such that computations can be performed with it. The basic assumption in the

---

B. Siddique (✉) · M. M. Sufyan Beg  
Aligarh Muslim University, Aligarh, India



**Fig. 1** Meaning preciation in CTP



**Fig. 2** Conceptual structure of computing With words (CWW)

CWW machinery is that the meaning conveyed by a natural language statement can be precisiated by constraining the values of variables contained in it. This precisiated form is called the Generalized Constraint Representation (GCR). The concept of GCR is central to the process of meaning preciation (and is discussed in detail in Sect. 3). However, not all the NL statements can be meaning precisiated in meaning. The set of NL statements which could be meaning precisiated is a subset of natural language and is termed as the precisiated natural Language. The meaning preciation is carried out using Constrained-centered Semantics of Natural Language (CSNL). For PNL, the associated precisiated language is called Generalized Constraint Language (GCL) [10]. Figure 1 depicts the discussion just presented.

Based on the concept of a generalized constraint, the high-level flow of CWW machinery is illustrated in Fig. 2. The initial data set (IDS) consisting of a set of natural language propositions is worked upon. The primary step is to carry out the meaning preciation by explicating the implicit semantics. This results into a set of constraint representations out of the IDS. Now, propagation is carried out based on the rules of inference in fuzzy logic to identify the output constraint. These identified constraints are then retranslated in the form of natural language propositions which constitute the terminal data set (TDS).

One of the primary concerns in the working of a CWW machinery (refer Fig. 2) is to deal with constraint explication. This involves the identification as well as the representation of the implicit constraints. In this paper, we deal with this subject for the lexical category of adjectives and adverbs. We have specifically addressed these lexical categories owing to the high information content carried by adjective phrases in scenarios where human perceptions are involved. Example of such scenarios include handling customer reviews, survey responses, comments to online media posts, etc.

The proposed generalized constraint representation in this paper serves as the formalization for conversion of the NL propositions containing adjective phrases into

their respective precisiated form. The study conducted finds application in solving complex CWW problems. In the literature, we find studies [4] on the similar subject. Following the proposal in the studies mentioned, we present concrete mathematical representations for phrases with single or multiple adjectives. Additionally, in this paper, we discuss the semantics of adjectives in the context of fuzzy variables.

The paper is organized as follows: In Sect. 2, we discuss the role and analysis of adjectives in English language. In Sect. 3, we brief the concept of a generalized constraint. In Sect. 4, we present the semantics of adjectives in light of fuzzy concepts. In Sect. 5, we elaborate our proposal of the GCR of adjectives. In Sect. 6, we report the implementation details. And finally, we conclude the paper.

## 2 Role and Analysis of Adjectives in the English Language

An adjective is defined as ‘A word naming an attribute of a noun, such as sweet, red, or technical [1]’. The attributes of the noun could be size, shape, age, color, origin, material, opinion, quality, type, purpose and participle forms. A sentence may contain multiple adjectives, the category of adjectives being same or different. When the number of adjectives in a sentence is multiple and belonging to different attributes, the adjectives usually appear in the following order [3]: opinion, size, quality, age, shape, color, participle forms, origin, material, type and purpose. Irrespective of the count or gender, an adjective appears in the same form in sentences. For example, the adjective *young* takes the same form in the phrases: a *young* boy, a *young* girl and a *young* generation. However, some adjectives take comparatives and superlatives endings. For example, Sam is *younger* than Bob, and Alice is the *youngest* among her siblings.

Depending upon the position in a sentence, adjectives could be either attributive or predicative. While attributive adjectives appear before the noun, predicative adjectives appear as a complement after a linking verb such as be, get and seem. For example, the adjective *red* is attributive in the sentence: It is a *red* car, whereas it is predicative in the sentence: The car is *red*. Notably, while most of the adjectives can occur at both positions, few can occur only in attributive or predicative positions only.

## 3 Concept of Generalized Constraint

The concept of Generalized Constraint Representation (GCR) is central to computing with words computation. The GCR of a NL proposition conveys the information contained in it. The basic assumption is that the NL proposition is assumed to be an answer to implicit question. The GCR reveals three parameters of a NL proposition as follows:

- The variable that is constrained by the proposition, termed as the constrained variable and represented by  $X$
- The value of the variable, termed as the constraining relation and represented by  $R$ , and
- The way in which the variable is constrained, termed as the modality of the constraint and represented by  $r$

Hence, for a proposition  $p$ , the GCR is represented as follows:

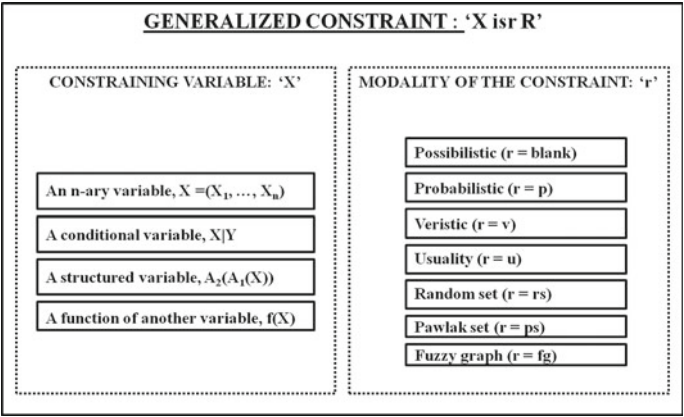
$$GCR(p) = X \text{ is } r R \tag{1}$$

Specifically,  $R$  depicts the label of the fuzzy set corresponding to the constraining relation, and  $X$  and  $r$  can take several forms some of which are listed in Fig. 3.

Consider, for example, the sample proposition *John is tall*. Assuming that this proposition is an answer to the implicit question: What is the height of John?, its GCR could be correspondingly represented as follows:

$$\text{height}(\text{John}) \text{ is } \text{tall} \tag{2}$$

where the constrained variable is the height of John, the constraining relation is the fuzzy set labeled *tall*, and the modality of the constraint is possibilistic.



**Fig. 3** Representation of entities in precisiation

## 4 The Semantics of Adjectives

### 4.1 Meaning Precision

In the context of meaning precision in Computational Theory of Precision of Meaning (CPTM), precision is categorized as either v-precise or m-precise [11]. While v-precise denotes precision in value, m-precise denotes precision in meaning. For example,

- X is 3: is v-precise as well as m-precise
- X is between 2 and 5: is m-precise but not v-precise
- X is small: is v-imprecise as well as m-imprecise, but it can be m-precised by defining a fuzzy set labeled ‘small’.

Based on this categorization, the lexical category of adjectives is characterized as v-imprecise as well as m-imprecise. Examples include handsome, beautiful, cunning, kind, etc. Moreover, all those adjectives which are associated with attributes which could be viewed as numerical variables (such as height, speed) could be made m-precise. This could be carried out by defining a corresponding fuzzy set. Examples of such adjectives include young, tall, fast, etc. The scope of this paper deals with m-precisable adjectives.

### 4.2 Semantics

In line with Zadeh’s concept of a linguistic variables [5–7], the semantics of adjectives is described as follows: Let the adjectives  $A_1, A_2, \dots, A_n$  describe the attribute *Att* of the noun *N*. Then, the set of adjectives  $A_1, A_2, \dots, A_n$  corresponds to fuzzy variables which constitute the values of the linguistic variable *Att*. Consider the attribute *age* for which a corresponding adjective could be *young*. Semantically, the fuzzy variable *YOUNG* depicts one of the values of the linguistic variable *AGE*.

On the basis of the formal definition of a linguistic variable, typically, the set of adjectives for an attribute comprise a few primary terms. Particularly, these primary terms are two in count and are antonyms to each other, with which other linguistic values are formed through a set of modifiers [8]. As an example, for the linguistic variable *AGE*, the primary terms include the adjectives *old* and *young*, and the set of modifiers could be consisting of terms like *too*, *very* and *not*. The computations for the composite terms could be carried out accordingly. For instance, supposing that the membership function for the fuzzy variable *OLD* is defined, then the possibility distribution for the proposition: *John is not very old* could be computed to be equal to  $(OLD^2)'$ , where the squaring operation approximates to the modifier *very* and the negation operation corresponds to the term *not*.



## 5 Generalized Constraint Representation of Adjectives

As discussed in Sect. 2, adjectives define noun attributes. Let the ‘attribute’ of a noun  $N$  be termed as the ‘characteristic function of  $N$ ’ and represented as  $CF(N)$ . Then, in line with the discussion of the Generalized Constraint Representation (GCR) in Sect. 3, for an adjective–noun pair:

- The constrained variable  $X$  is the characteristic function of the noun which the adjective denotes,
- The constraining relation  $R$  is the fuzzy set depicted by the adjective/s (details are mentioned in following subsections for single and multiple adjective cases)
- The modality of the constraint,  $r$ , is possibilistic.

Accordingly, the GCR of an adjective–noun pair follows:

$$f(X) \text{ is } R \rightarrow CF(N) \text{ is adjective} \quad (3)$$

Correspondingly, the protoform for the lexical category of adjectives is  $B(A) \text{ is } C$ , where  $A$ ,  $B$  and  $C$  are the abstractions for the noun, characteristic function of the noun depicted by the adjective and the adjective, respectively. Following, we present the mathematical interpretation of the proposed constraint representation for the two cases corresponding to noun phrases consisting of single adjective and multiple adjectives.

### 5.1 Noun Phrases/ NL Propositions Consisting of Single Adjective

If a noun phrase contains an adjective  $\langle Adj \rangle$  attached to the noun ‘ $X$ ’ such that the  $\langle Adj \rangle$  determines the characteristic ‘ $f$ ’ of the noun ‘ $X$ ’ to be ‘ $Y$ ’, then the noun phrase could be represented in the form of a generalized constraint given as follows:

$$f(X) \text{ is } Y \quad (4)$$

where  $Y$  is the label of a fuzzy set with universe of discourse  $U$  and membership function  $\mu_Y$ . Applying the theory of possibility [9], the possibility distribution induced by the adjective on the noun’s attribute is given as follows:

$$\Pi_{f(X)} = Y \quad (5)$$

such that the possibility of the noun’s attribute to take its value  $u$  in  $U$  is given as  $\mu_Y(u)$ .

**Example:** For the proposition ‘Monika is young’ consisting of single adjective *young*, the generalized constraint representation is

$$Age(Monika) \text{ is } young \quad (6)$$

where *young* is the label of the fuzzy set with universe of discourse, say,  $U = [0, 100]$  and a membership function  $\mu_{young}$ . The induced possibility distribution is

$$\Pi_{Age(Monika)} = young \quad (7)$$

## 5.2 Noun Phrases/ NL Propositions Consisting of Multiple Adjectives

If a noun phrase contains  $k (> 1)$  adjectives denoted as  $\langle Adj_i \rangle$  ( $i = 1$  to  $k$ ) attached to the noun ‘X’, and  $\langle Adj_i \rangle$  determines the characteristic ‘ $f_i$ ’ of the noun ‘X’ to be ‘ $Y_i$ ’, then the noun phrase could be represented in the form of a generalized constraint given as follows:

$$f_1(X), f_2(X), \dots, f_n(X) \text{ is } Y \quad (8)$$

where Y is the label of a fuzzy set on  $n$ -tuples with universe of discourse  $U_1, U_2, \dots, U_n$  corresponding to  $Y_1, Y_2, \dots, Y_n$ , respectively, and membership function  $\mu_Y$ . In the theory of possibility [9], Y is an  $n$ -ary possibility distribution corresponding to the fuzzy relation in the Cartesian product  $U_1 * U_2 * \dots * U_n$  is given as follows:

$$\Pi_{f_1(X), f_2(X), \dots, f_n(X)} = Y \quad (9)$$

such that the possibility of the noun’s attributes to take its value  $(u_1, u_2, \dots, u_n)$  is given as  $\mu_Y(u_1, u_2, \dots, u_n)$ . Assuming the attributes of the noun to be non-interactive, the induced possibility distribution reduces to

$$\Pi_{f_1(X), f_2(X), \dots, f_n(X)} = \Pi_{f_1(X)} * \Pi_{f_2(X)} * \dots * \Pi_{f_n(X)} \quad (10)$$

**Example:** For the proposition ‘Young boys are quick’ consisting of two adjectives, *young* and *quick*, the generalized constraint representation is

$$Speed(boys), Age(boys) \text{ is } Y \quad (11)$$

where Y is the label of 2-tuple fuzzy set with universe of discourse, say,  $U_{age} = [0, 100]$  and  $U_{speed} = [0, 10]$  and a membership function  $\mu_Y$ . Assuming *age* and

*speed* to be non-interactive, the induced possibility distribution is the fuzzy relation in the product of *quick* \* *young* given as follows:

$$\Pi_{Speed(boys), Age(boys)} = \Pi_{Speed(boys)} * \Pi_{Age(boys)} \quad (12)$$

## 6 Implementation Details

For the conversion of NL propositions containing adjectives to generalized constraints, we undertake the following steps:

1. Parsing the NL proposition to identify adjective/s and nouns: For the purpose, we have used the Stanford Parser [2]. The relations ‘amod’ and ‘nsubj’ are used to identify adjectives at attributive and predicative positions, respectively.
2. Identification of the characteristic function/attribute name of the adjectives: For this, we have manually constructed a concept lookup table which stores adjectives in the form of categories characterized by the attribute which they depict. Some of the attributes of the concept lookup table include the ones mentioned in Table 1. In the implementation, after identifying an adjective as the first step, a lookup is made in the concept lookup table to assign the corresponding attribute name to it. Following this, the text constituting the proposed representation is built as the final output.

The current implementation falls short in handling polysemous adjectives, i.e., those adjectives which depict different meaning in different contexts. As an example, consider the role of adjective *short* in the following propositions:

- The journey is *short*.
- Simon is *short*.

While in the former, *short* depicts the ‘distance’ attribute of the noun journey, in the latter, it depicts the ‘height’ attribute of the noun Simon. The current implementation does not take into account the context, and the attribute assignment depends on the match which is found first in the concept lookup table.

## 7 Conclusions

Zadeh’s paradigm of Computational Theory of Perceptions (CTP) serves to deal with NL-based information. This paper addressed the subject of constraint explication in the framework of PNL. Specifically, the lexical category of adjectives is dealt with owing to the fact that it contains significant amount of information in human perception-based data. We present the semantics of adjectives followed by the proposal and implementation for formalizing its Generalized Constraint Representation

(GCR). The proposed GCR for adjectives is  $f(X)$  is  $R$ . The scope of this paper deals with m-precisable adjectives. We intend to deal with m-imprecise adjectives in the future.

## References

1. Definition of adjective in English by Oxford Dictionaries. URL <https://en.oxforddictionaries.com/definition/adjective>
2. Stanford CoreNLP. URL <https://stanfordnlp.github.io/CoreNLP/index.html>
3. Eastwood J (1994) Oxford guide to English grammar. Oxford University Press
4. Soto A, Olivas JA, Prieto ME (2007) Using generalized constraints and protoforms to deal with adjectives. In: Fuzzy Systems Conference, FUZZ-IEEE 2007. IEEE International, pp 1–6. IEEE
5. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-iii. Information Sci 9(1):43–80
6. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-i. Information Sci 8(3):199–249
7. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-ii. Information Sci 8(4):301–357
8. Zadeh LA (1978) PRUF—a language for the representation of meaning in natural languages. Int J Man-Machine Stud 10(4):395–460
9. Zadeh LA (1999) Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst 100(1):9–34
10. Zadeh LA (2001) A new direction in ai: toward a computational theory of perceptions. AI Mag 22(1):73–84
11. Zadeh LA (2006) From search engines to question answering systems-the problems of world knowledge, relevance, deduction and precisiation. In: Capturing intelligence, vol 1, Elsevier, pp 163–210

# ECG Signal Compression Using Singular Value Decomposition and Implementation on PYNQ-Z2



Mayur M. Sevak and Tanmay D. Pawar

**Abstract** In today's daily lifestyle of human beings, most people are more aware of individual health, and also, they are using smart wearables with various features that they can measure easily and have statistics of it. An electrocardiogram measures the electrical activity of the heart. For cardiac patients' diagnosis, one has to go for long-term monitoring of ECG signal which also leads to deal with a huge amount of data that finds its applications for signal compression. In this work, the authors have taken ECG signals from various persons through BIOPAC systems with ECG lead-I configuration. All ECG data have been compressed by Singular Value Decomposition on the PYNQ-Z2 FPGA where it provides a programming platform on python. SVD is a matrix-based method where the signal is decomposed into a set of basic patterns and scaling factors. The most interesting thing about SVD is that very few singular values can retain most of the information of the signal. Here we have recorded ECG signals with BIOPAC systems with a sampling frequency of 1 K S/s and After applying SVD to them we have reconstructed signals of sample sizes 20,000 and 7000. The maximum Compression ratio that we achieved is 37.9506 and minimum PRD(%) 1.4797 which is same as PRRE(%) 1.4797 for 20,000 samples and The maximum Compression ratio that we achieved is 13.7254 with minimum PRD(%) 1.3077 which is same as PRRE(%) 1.3077 for 7000 samples.

**Keywords** Electrocardiogram · Data compression · Singular value decomposition · PYNQ-Z2 · Application processing unit (APU) · Programmable logic structure · Compression ratio · Percent root mean square difference (PRD) · Percent root mean square residual energy (PRRE)

---

M. M. Sevak (✉)  
Gujarat Technological University, Gujarat, India  
e-mail: [mayur.sevak@bvmengineering.ac.in](mailto:mayur.sevak@bvmengineering.ac.in)

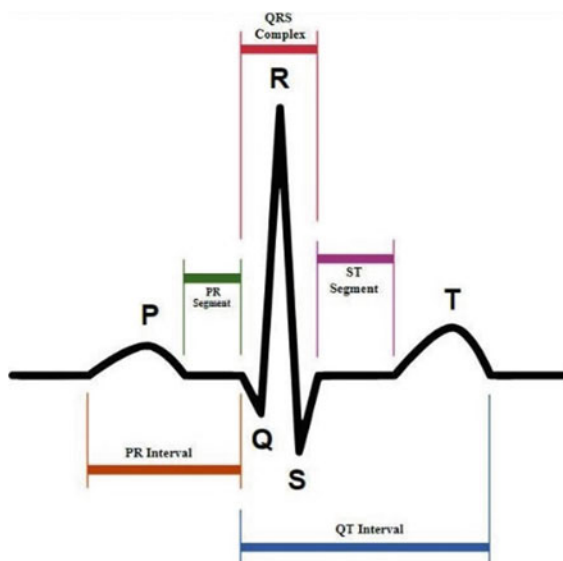
T. D. Pawar  
Electronics Engineering, Birla Vishvakarma Mahavidyalaya, Anand, Gujarat, India  
e-mail: [tdpawar@bvmengineering.ac.in](mailto:tdpawar@bvmengineering.ac.in)

# 1 Introduction

Sudden Cardiac Attack is considered to be a fast-growing disease in the twenty-first century for persons of any age. This has led to measure the health of the heart. ECG is the electrical activity of the heart which has various features like *P*-Wave, PR interval, QRS complex, J point, ST segment, *T* wave, QTc interval, *U* wave. The normal cardiac cycle starts with the sinus node's spontaneous depolarization which is an area of specialized tissue located in the high Right Atrium(RA). The depolarized electrical wave spreads through RA and across the internal atrium septum into Left Atrium. The inert fibrous ring electrically separates atria from ventricles so that in a normal heart, electrical depolarization finds its route for transmission from atria to ventricles through the atrioventricular(SA) node. The AV node is responsible for delaying electrical signal for short time and then a wave of depolarization spreads down the interventricular septum(IVS) via the bundle of His and right and left bundle branches into right and left ventricles. Hence two ventricle contracts simultaneously in normal conduction. After complete depolarization of the heart, the myocardium must then repolarize-Wave is formed due to depolarization of atrial musculature. The QRS complex is formed due to depolarization of ventricular musculature wave is form due to repolarization of ventricular musculature (Fig. 1).

For long-term monitoring of ECG data or in an ambulatory monitoring system we need to analyze a huge amount of data to get enough information about the person under-diagnosis. This led to compress the ECG signals effectively for many applications like transmission of ECG signals over telecommunication networks, ambulatory ECG recorders, ECG data storage. Even a real-time monitoring system

**Fig. 1** ECG wave [1]



of ECG requires real-time data compression and transmission which uses wired or wireless communication resources. ECG signals of cardiac patients need to be stored for monitoring as well as diagnosis purpose. Since this is bulky so it is required to make it smaller and compatible for storage and transmission purpose for real-time applications. The objective of any data compression technique is to reduce the information size by removing the redundancy in the signal while keeping diagnostic features of the signal. Data compression techniques are classified into the following categories: parameter extraction techniques (average beat subtraction, cycle pool-based subtraction [2], Direct data compression technique (DPCM, AZTEC, SAPA), Transform methods (KL transform, Wavelet transform) it is also classified as lossy as well as lossless compression.

Here, in this work, we have taken ECG signals from different persons having different ages. Their ECGs have been taken through BIOPAC data systems in Lead I configuration with a sampling frequency of 1 kHz. After that, all the ECG signals were filtered and peak detection has been performed with the first derivative method for QRS Complex detection [3] followed by Beat Alignment of all ECG beats. SVD has been applied on this matrix in which few numbers of singular values are kept and redundancy is removed. Performance is measured through compression ratio.

## 2 Literature Review

The authors Anlikumar et al. [4] discussed SVD computation with Jacobi Method, QR Algorithm, and one-sided Hestenes Jacobi method where they found that the Jacobi method has higher accuracy for singular values than the QR algorithm for parallel implementation [4]. In that Jacobi algorithm, BLV proposed an array architecture that can compute the SVD of  $n \times n$  real matrix. The hardware implementation of the Hestenes Jacobi method consists of ROM of  $(m \times n)$  matrix, two-port RAM, Control unit, Hestenes Jacobi block, Jacobi rotation block for orthogonalization of rows and elimination of off-diagonal elements, Square root block for calculating singular values were utilized. SVD for ECG compression was done by Syed Salman Kabir et al. [2] where they used data extraction and TSVD for ECG signal compression. They had taken amplitude difference and then compared with threshold 0.001 and 0.005 followed by SVD. Results show that for threshold value 0.001 and singular values 2 CR = 10.59 is achieved and for threshold value 0.005 and singular values and singular values 2 CR = 11.12 is achieved. Aidin et al. [5] had implemented singular value decomposition on FPGA where authors had customized the hardware structure by using the Jacobi algorithm for FPGA for computing SVD matrices. The reason for the selection of the Jacobi algorithm was because of the high parallel processing potential. They have implemented new method for determining the number of sweeps with the CORDIC algorithm to calculate basic trigonometric functions that had resulted in improvement of hardware and calculation speed. Ting et al. [6] had done ECG data encryption and compression using SVD where they got resultant matrix  $C$  by applying key matrix  $Q$  on ECG data of  $Y$  matrix. They

had taken 10 min ECG records of the MIT-BIH database with a sampling frequency of 360 Hz and 11 bits/sample resolution. They achieved  $CR = 40.90$  for singular value  $Q = 5$  for non-encrypted data  $CR = 39.8$  for 38.9 for encrypted data. With  $Q = 13$  algorithm was compared with other 12 compression technique and achieved  $CR = 50.34$ . Dargie et al. [7] had utilized SVD to remove motion artifacts from the measurement of wireless ECG where they used 5-lead ECG, 3D-accelerometer, 3D-gyroscope, 3D-magnetometer with various activities and movements. They had got a result that SVD was more efficient and consistent with the detection of QRS complex in all movements and found reliable for detection of  $P$  and  $T$  waves during regular movements.

### 3 Proposed Model

#### 3.1 Software Platform

##### 3.1.1 Matlab 2019

For ECG detection of QRS complex and ECG beats alignment. Following is the method used for the detection of the QRS Complex.

##### 3.1.2 First Derivative Method for QRS Detection

The first derivative  $y(n)$  is calculated for each point of ECG signal of  $x(n)$  mentioned in [3] as:

$$Y(n) = [X(n+2) - 2X(n-2) + X(n+1) - X(n-1)] \quad (1)$$

The slope threshold ( $Th$ ) is calculated using  $0.6 * \max(Y)$  [3]. Searching was applied to the first derivative-based array for points that exceed the slope overload. Now the points that satisfy the condition are taken as QRS and their location was marked to generate the ECG Beat matrix where each column of the matrix is a single beat of ECG.

##### 3.1.3 Jupyter Notebook Python 3

Here we have used following python libraries for implementation of SVD.

1. Pandas: it is used to load the ECG signal.
2. Matplotlib: For plotting the ECG Beat Matrix and Reconstructed ECG Beat Matrix.
3. Heartpy: For preprocessing of the ECG Signal



4. SciPy: For SVD function implementation.

### 3.1.4 Singular Value Decomposition

Let's consider an Electrocardiogram Signal with  $p$  periods and  $q$  length. Each period is an  $R$ - $R$  interval. This signal can be arranged in a two-dimensional matrix having  $p \times q$  dimensionality. Each row of the matrix shows each period of ECG signal [2] shown in Eq. (2).

$$M = \{m_i(t) | i = 1, 2, 3, \dots, p, t = 1, 2, 3, \dots, q\} \quad (2)$$

$$\begin{pmatrix} m(1) & m(1) & m(q) \\ m(q+1) & m(q+2) & m(2q) \\ \vdots & \vdots & \vdots \\ m((p-1)q+1) & m((p-1)q+2) & m(pq) \end{pmatrix}$$

where  $m_i(t)$  is  $i$ th period of  $m(t)$ . SVD of given matrix having dimension of  $p \times q$  can be found by [2] as shown in Eq. 3 given below.

$$M = U \sum V^T \quad (3)$$

where  $U \in R^{p \times p}$ ,  $V \in R^{q \times q}$  are left and right singular vectors.  $\sum$  is  $p \times q$  matrix where  $X_1, X_2, X_3, \dots, X_n$  are singular values.

Singular Value Decomposition theorem says  $M$  can be decomposed into scaling factors  $u_i$  and  $v$  vector where  $i$  is ranging from 1 to  $R$ . if matrix  $M$  is having rank one then except  $\times 1$ , other singular values will be zero. To reconstruct that matrix again, only principal components  $v_1$  and  $u_1 \times 1$  are required.

Hence, from [2] reconstruction can be implemented by Eq. (4).

$$M' = u_1 x_1 v_1^T \quad (4)$$

Now, ECG is a periodic signal with characteristics as compared to the sinusoidal signal in that it is containing periodic as well as aperiodic signals, and with that  $M$  can be the full rank matrix.

The information energy can be shown in terms of [2] given in Eq. 5.

$$Q_M = x_1^2 + x_2^2 + \dots + x_R^2 \quad (5)$$

and main information we can still recovered from very few singular values if [2] ratio is given in Eq. (6)

$$x_1^2 / x_i^2 \gg 1 \quad (6)$$

where  $i = 1, 2, 3, 4, \dots, N$ .

from the above  $N$  singular values, if only  $r$  values are greater than the remaining values then we can remove those nonrequired values from  $\sum$  and matrix  $m$  can be reconstructed by [2] with the equation given (7) given below.

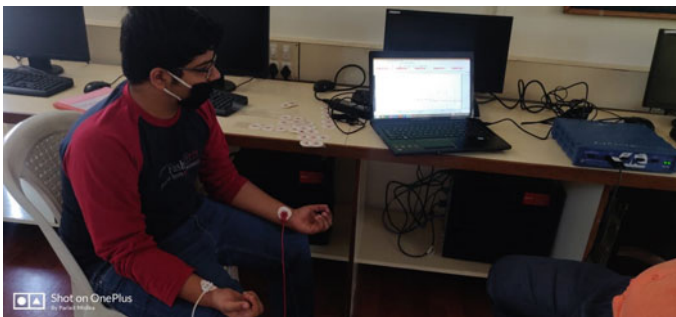
$$M' = u' \sum' v'^T \quad (7)$$

where  $U' \in R^{p \times p}$ ,  $\sum' \in R^{r \times r}$ ,  $V' \in R^{q \times q}$  and  $U'$  and  $V'$  are having orthonormal columns.

## 3.2 Hardware Platform

### 3.2.1 Biopac Student Lab System Using MP Acquisition

MP data acquisition unit of the Biopac Student Lab PRO system is having a pivotal role in the data acquisition system. It has an internal microprocessor that controls data acquisition and it can communicate with the computer. It is having analog channels four on MP3X units and two on MP45 units). In this research, we have used MP36 which is having a resistance range of 0–1 M $\Omega$  for electrode checking, 4 isolated channels available for analog inputs. it is capable of providing 100 K s/second. it provides 24-bit A/D resolution, signal to noise ratio > 89 dB. it is having voltage resolution of 2.38 microvolts/bit (Gain 5) to 0.024 nanovolts/bit (Gain 50,000). The maximum input voltage that it can support is 4 V  $p-p$ . it provides 3 two-pole IIR digital filters per channel, CMRR 110 dB at 50/60 Hz. It provides 1 channel for analog output and 16 bits resolution for D/A. For serial communication, it is having USB 2.0 and I/O ports that are 8 TTL compatible. Figure 2 shows the practical setup for taking ECG signals in lead I configuration through the Biopac Student Lab system



**Fig. 2** ECG acquisition of person in Lead I configuration using BSL

**Fig. 3** Multi lead ECG cable

with the sampling frequency of 1 K s/s and the output of the signal was exported via BSL software in.txt and.xlsx format.

Figure 3 shows the SS29 L multi-lead ECG signal that permits ECG recordings with high resolution. This set can perform the simultaneous recording of various leads like I, II, III, aVR,aVL,aVF along with one precordial chest lead V(1–6).12 leads ECG recording is also feasible with altering the chest leads. The cable terminates in three smart sensors connected to MP36.

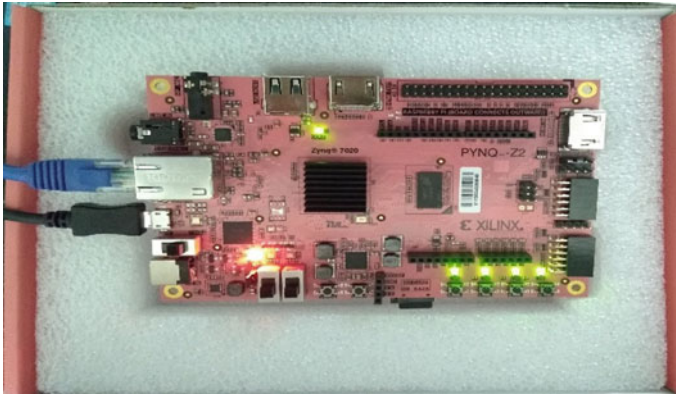
### 3.2.2 PYNQ-Z2 Development Board

PYNQ is an open-source platform provided by Xilinx to develop an embedded system with ease of use where it supports python programming and various libraries. Before the development of Zynq architecture, it was complicated to couple Field Programmable Gate Array and Microprocessor. Zynq architecture is the latest generation of Xilinx's SOC family which combines dual-core ARM Cortex-A9 with FPGA. Hence it is having a Processing system(PS) and Programmable Logic(PL). PYNQ-Z2 board is shown in Fig. 4.

- Application Processing Unit:

Application processing unit of PYNQ-Z2 consists of two ARM cortex-A9 processors where each unit includes following.

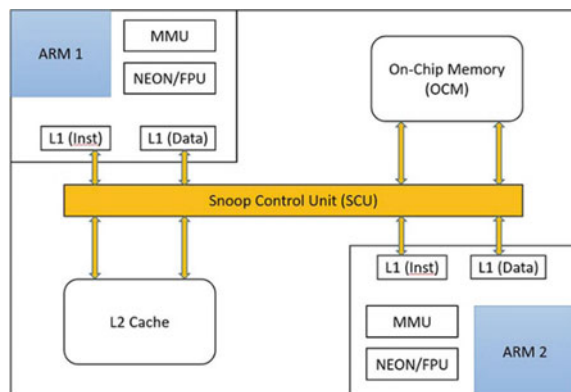
- Neon: This unit provide Single Instruction Multiple Data which accelerates DSP, Media algorithm to an ARM processor.
- FPU: it gives acceleration to floating point operations.
- Level 1 cache: Every processor has its own data and instruction caches for storing them.
- MMU: it translates virtual memory addresses to physical memory addresses.
- Snoop control unit: it provides interfacing between processors, L1 and L2 caches.



**Fig. 4** PYNQ-Z2 development board

- L2 cache: it is shared between the two processor that enables the newest update of variable all are shown in given Fig. 5.
- Programmable Logic Structure:
  - It consists of configurable logic blocks with two slices. Each is containing four look-up tables, eight flip-flops, a switch matrix, Block RAM and DSP slices.
  - Slice: Each slice is having resources to implement combinational and sequential circuits.
  - LUT: To implement logic function up to six inputs, ROM, RAM, or shift registers
  - FF: To implement a sequential element of 1-bit register with reset.
  - Switch matrix: it connects different parts within and between CLB and other parts of PL (Fig. 6).
- Features of PYNQ-Z2:

**Fig. 5** Structure of application processing unit [8]



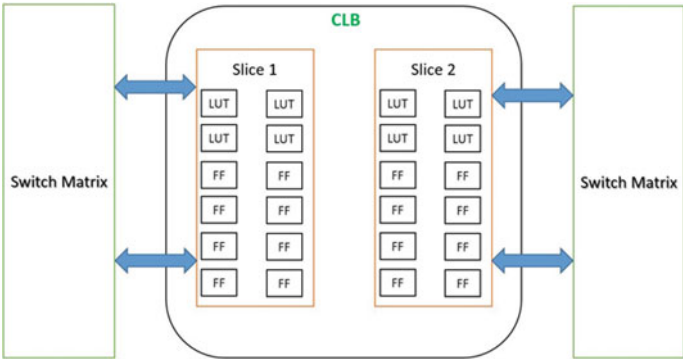


Fig. 6 PL structure [8, 9]

It is having a 650 MHz ARM Cortex A9 dual processor,13,300 logic slices each with 6-input LUT and 8 flip-flops,630 KB Block RAM,220 DSP Slices-ADC, Programming with JTAG, MicroSD card, Quad-SPI Flash. Memory and Storage of 512 MB DDR3 with the 16-bit bus which supports 1050 Mbps,16 MB Quad-SPI Flash. Micro SD slot. It supports USB and Ethernet for programming and other purposes.

4 Block Diagram and Flow of Our Model

Figure 7 shows the block diagram and flow of work. In this, we have taken the ECG signals of four persons having different ages. Their ECGs were recorded in the lead 1 configuration as shown in Fig. 2 using BIOPAC systems. After that applying the filter, the baseline wander was removed. By using MATLAB 2019, peak detection mentioned in Sect. 3.1.2 was applied to the signal. First, we took 20,000 samples

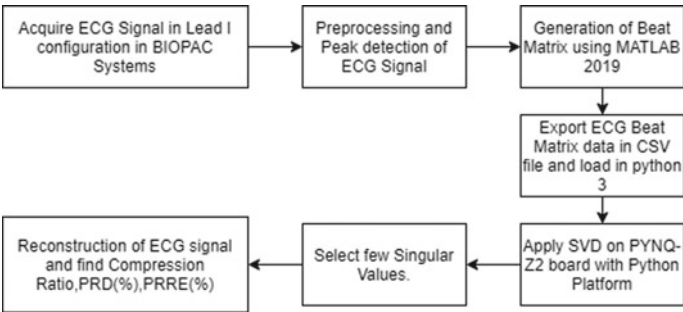


Fig. 7 Block diagram and flow of our model

**Fig. 8** Implementation of SVD on PYNQ-Z2



of ECG and generate ECG Beat Matrix. This ECG Beat matrix data was loaded on the PYNQ-Z2 python environment where it uses the PS part of Zynq architecture. SVD was applied to the signal. Reconstruction of ECG was carried out with a smaller number of singular values and Compression Ratios were calculated. The same exercises were repeated for 7000 samples as well.

Figure 8 shows an implementation of SVD on ECG Beat Matrix where we have used PS of PYNQ-Z2 board. Initially, 20,000 samples of ECG signals were taken and we have applied SVD with Python Programming. Reconstruction of ECG Signal was done with few numbers of singular values. The same exercises were carried out for 7000 ECG Samples and the compression ratio was found for the same.

## 5 Performance Parameters and Simulation Results

### 5.1 Compression Ratio

$$CR = \frac{b_0 \sum_{i=1}^P T_i}{(M + N + 1) \times q \times b_s + P x b_p + (b_\alpha + b_\beta + b_\gamma) \times 2} \quad (8)$$

here, in the numerator  $b_0$  represents quantization level bits/sample which is multiplied with the summation of all ECG period length.  $P$  is the total number of ECG segments and  $T_i$  represents  $i$ th ECG cycle. In the denominator,  $M$  represents the number of elements of the left singular vector( $u_i$ ) which is the same as  $P$ .  $N$  represents the number of elements of the right singular vector( $v_i$ ),  $q$  is truncated index.  $b_s$  represents the quantization level of singular triplets. The second term is utilized for the storage of associated beat information and represents bits for storage of each segment. Final terms consist of miscellaneous parameters including  $b_\alpha$ ,  $b_\beta$ ,  $b_\gamma$  which are showing the number of bits to store the first singular value, number of beats, length of shortest and longest period.

**Table 1** SVD and Reconstruction of various person's ECG signals with sample size = 20,000, sampling freq = 1 k s/s

Subject	No of singular values	Compression ratio	PRD(%)	PRRE(%)
Person 1	3	18.0342	1.9450	1.9450
	2	24.3013	2.4181	2.4181
	1	37.2439	4.6446	4.6446
Person 2	3	19.9203	3.2078	3.2078
	2	24.6609	3.5959	3.5959
	1	37.6647	5.8898	5.8898
Person 3	3	18.4842	9.9179	9.9179
	2	24.8447	11.1612	11.1612
	1	37.8787	14.7979	14.7979
Person 4	3	18.5356	1.4797	1.4797
	2	24.9066	1.9436	1.9436
	1	37.9506	4.1794	4.1794

## 5.2 Percent Root Mean Square Difference

$$\text{PRD}(\%) = \sqrt{\frac{\sum_{i=1}^L [(x_0(i) - x_r(i))^2]}{\sum_{i=1}^L x_0^2(i)}} \times 100 \quad (9)$$

The percent root mean square difference is parameter that shows the error between original data and reconstructed data.

## 5.3 Percent Root Mean Square Residual Energy

$$\text{PRRE}(\%) = \sqrt{\frac{\sum_{i=1}^L [x_0^2(i) - x_r^2(i)]}{\sum_{i=1}^L x_0^2(i)}} \times 100 \quad (10)$$

The percent root mean square residual energy is the energy information of reconstructed signal and it is evaluating factor to determine reconstructed error. Since  $x_0(i)$  is very close to  $x_r(i)$  (Tables 1 and 2; Figs. 9 and 10).

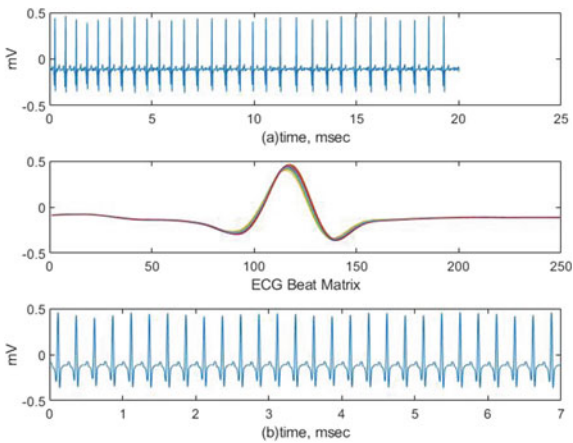
## 6 Conclusion

In this work, we have recorded ECG signals of 3 min duration from various persons with BIOPAC systems. After applying SVD to all ECG Signals on the PYNQ-Z2

**Table 2** SVD and reconstruction of various person’s ECG signals with sample size = 7000, sampling freq = 1 K S/s

Subject	No of singular values	Compression ratio	PRD (%)	PRRE (%)
Person 1	3	6.7307	1.4408	1.4408
	2	9.009	1.7840	1.7840
	1	13.618	3.7578	3.7578
Person 2	3	6.750	2.8866	2.8866
	2	9.3022	3.4900	3.4900
	1	13.6452	6.2966	6.2966
Person 3	3	6.8093	7.7187	7.7187
	2	9.1027	10.0950	10.0950
	1	13.7254	15.4497	15.4497
Person 4	3	6.7895	1.3077	1.3077
	2	9.079	1.7048	1.7048
	1	13.6986	3.3794	3.3794

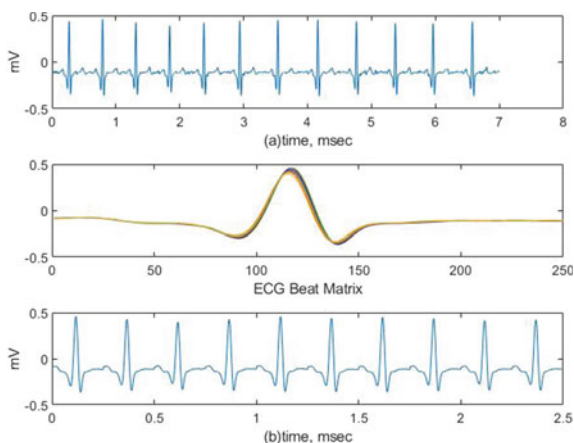
**Fig. 9** Input ECG, beat matrix and reconstructed ECG of person2 with singular value = 2 and sample size = 20,000 [9]



platform (PS), we can reconstruct ECG signals with very few singular values. PYNQ is providing a flexible environment for coding in python hence it was easy to apply SVD on ECG Beat Matrix. Experimental results show that dimensionality reduction is achieved were for 20,000 samples of ECG we got maximum Compression Ratio 37.9506 and minimum PRD(%) 1.4797 which is same as PRRE(%) 1.4797 and for 7000 samples of ECG we got maximum Compression Ratio 13.7254 with minimum PRD(%) 1.3077 which is same as PRRE(%) 1.3077. It is also observed that as number of singular value increases then value of CR is decreased and PRD and PRRE decreased.



**Fig. 10** Input ECG, beat matrix and reconstructed ECG of person 1 with singular value = 2 and sample size = 7000



## References

1. Jyh-JongW, Chuang-Jan C, Nai-Kaun C, Gwo J (1997) ECG data compression using truncated singular value decomposition. *IEEE Trans Biomed Eng* 44(1). <https://www.alivecor.com/education/ecg.html>
2. Syed K, Mamshad R, Kamrul HM, Sneddon I (2017) ECG singnal compresion using data extraction truncated singular value decomposition. In: *IEEE region 10 humantarian technology conference*, Dec 2017
3. Kher R,Vala D, Pawar T, Thakar V (2010) Implementation of derivative based QRS complex detection methods. In: *IEEE international conference on biomedical engineering and informatics*
4. Anilkumar S, Swanirbhar M, Subir S (2013) Hardware implementation of singular value decomposition. In: *Proceedings of Michel Faraday IET India summit*
5. Aidin S, Karamian G (2019) An FPGA implementation of singular value decomposition. In: *27th Iranian conference on electrical engineering*
6. Ting L, Kuan L, Wu H (2018) ECG data encryption then compression using singular value decomposition. *IEEE J Biomed Health Inf*
7. Waltenegus D, Jannis L (2019) Application of SVD for removing motion artifacts from measurments of wireless electrocardiogram. In: *IEEE international conference on information fusion*
8. Abenstein J, Tompkins W (1982) A new data reduction algorithm for real time ECG analysis. *IEEE Trans Biomed Eng* **BME-29**(1)
9. [www.aldec.com](http://www.aldec.com)

# Multi-Sectoral ICT Applications Toward Making Sustainable Rural India



Sachi Gawai, Jyotshna Dongardive, and Pallavi Chatuphale

**Abstract** An Information and Communication Technology (ICT) in agriculture is an arising field zeroing in on the upgrade of rural area in India. It includes utilization of inventive approaches to utilize ICT in the country space. It can give exact data essential for the farmers which works with better agricultural yield. In ongoing year, farmers demeanor to admittance agricultural data have been changed on account of quick systems administration of data and correspondence innovation. Farmers can get the data with respect to composts, pesticides, crop examples and climate estimating and other data through zero reasonable expense or minimal expense. A large number of the associations like government, private, co-agents, and public have likewise endeavored to work with the data innovation move in the agribusiness area. ICT is critical in working with correspondence and admittance to data for agricultural and country advancement. With this viewpoint, paper discovers the pertinent ICT applications for rural augmentation in India concerning public sector, non-governmental organization and private sector arrangements and projects viewpoint.

**Keywords** ICT · Agriculture · Private sector · Public sector · Non-governmental organization · Policy

## 1 Introduction

Agriculture is the essential wellspring of occupation for about 58% of India's population. India remains among top three as far as creation of different agrarian products like paddy, wheat, beats, groundnut, assault seeds, natural products, vegetables, sugarcane, tea, jute, cotton, tobacco leaves, and so on [1]. During 2019–20 harvest year, food grain creation arrived at a record of 296.65 million tons. In 2020–21,

---

S. Gawai (✉)

Ram Rao Adik Institute of Technology, University of Mumbai, Mumbai 400098, India

J. Dongardive · P. Chatuphale

University Department of Computer Sciences, University of Mumbai, Mumbai 400098, India

e-mail: [pachatuphale@pesbncollege.org.in](mailto:pachatuphale@pesbncollege.org.in)

Government of India is focusing on food grain creation of 298 million tons. Disregarding this impressive development, the enormous difficulties confronting Indian farming are to additional expansion the creation to stay up with the always in-wrinkling request from developing population. Indian agribusiness as far as economy is a long ways behind because of poor systems administration of market, suspicious information given to the farmers, land possessions, ranch produce value rates, and some more [2]. Information Communication Technology (ICT) in agriculture is being considered as a conceivably amazing passage point. In huge number of non-industrial nations, ICT in agriculture has been instrumental in changing poor agrarian practices, yields and incomes into cultivating that can financially maintain and improve the vocations of countless little farmers. To get the ideal outcomes from the utilization of ICT for dispersal of data in a country where larger part of the farmers are uneducated, land property are little or minor, and the degree of framework improvement is poor in the rustic regions, there is need to evaluate the data prerequisite of the farmers [3]. Further, how successfully ICT might be utilized to convey the necessary data as per the general inclination of the client and distinguishing the appropriate model for Indian farmers is required. Against this foundation, the current paper makes an endeavor to audit a portion of the ICT drives made under the government, non-governmental organization, and private sector.

## 2 Brief Overview

ICT has numerous jobs to perform for farming development. “E—Agriculture” is term utilized when ICT comes in enhancement of horticulture a provincial turn of events. Conceptualization, design, improvement, assessment, and use of creative ways are associated with ICT-based farming. In 2008, the United Nations alluded to e-farming as an arising field. ICT implies devices, networks, mobiles administrations, and application which go from imaginative web advancements and sensors to other prior assets like fixed phones, TV, radios, and satellite [4]. Examples of overcoming adversity on data and correspondence advances for horticulture and provincial improvement like in forecasting, seed, and trade community have archived numerous instances of utilization of ICT in agriculture [3, 5]. The different agro status areas, size of land property, types of yields developed, innovation followed, market direction, and climate conditions such data is needed in agriculture is done by ICT. It also helps farmers in choosing selection for jumping up with kind of yields, rehearsing shrewd farming practices for developing, reaping, post collecting and showcasing their demonstrate to actuate higher outcomes. Some devices of ICT are describe below.

- **Mobile App:** Mobile app facilitates a single window solution to the farmers and investors so as they can spread information about farm machinery, fertilizers, and soil health. It acts as bribed in connecting rural to urban India [6].

- **SMS Services:** mKisan SMS portal for farmers empowers all Central and State government organizations in agribusiness and partnered areas to give data/administrations/warnings to farmers by SMS in their language, inclination of agrarian practices and area [7].
- **Call Centers:** Queries identified with farming and associated areas are being tended to through these call places. A farmer from any piece of the State can contact the Kisan call center by dialing the complementary Telephone No. 1551 or 1800-180-1551 and present their issues/questions identified with cultivating [8].
- **E-Forecasting/News:** News provided by the government official helps in forecasting. It permits farmers to precisely explore to explicit areas in the field, after a seemingly endless amount of time after year, to gather soil tests or screen crop conditions. Area data is gathered by GPS beneficiaries for planning field limits, streets, water system frameworks, and trouble spots of harvests like weeds or sickness [9].
- **E-Learning/E- Marketing:** Vital data for agrarian advancement can be done on radio for instance data on better cultivating techniques, improved seeds, ideal planting, agro-ranger service, better gathering strategies, soil protection, showcasing, post collect taking care of and enhancement [1, 10].
- **Agriculture Web Portal:** Provides total data identified with farming for every area in region of each province of India, This agribusiness-related data incorporates crop filled around there, its seeds assortments, its market costs, hardware utilized, soil and climate advices, stockpiling, its sellers and various government plots, etc. [6].
- **Social Media:** Social media is now a mainstream form of communication around the world, and continues to grow in popularity with the increase in the number of smart phones. Twenty-first century has given the power to promote wellbeing of human nature in every aspect. Hence, hardship owed by farmers in agriculture can be expressed and spread their purpose through social media like Facebook, YouTube [11].

### 3 ICT Drives Under Public Sector

Agriculture is one of the prominent areas of Indian economy giving jobs to the greater part of India's population. While giving food security, this is the foundation of the country's rural economy. In the primary quarter of FY 2020–21, when the Indian economy enrolled 23.9% negative development, agribusiness was the lone area which arose as a silver fixing for India's monetary recuperation with over 3.4% development. It makes farming the most noticeable area in country. Furthermore, consequently, it needs a great deal of public sector support for a manageable development [12]. The salient features of these government schemes in agriculture sector are:

- To provide services on pesticide, manures, seeds, and fertilizers and insecticides
- Soil health management
- New innovative farm machinery techniques and practices.

- Information on weather forecasting and agro-met advisory
- Import and export informants
- Market strategies
- Evaluation of schemes and upcoming model programs
- Irrigation management
- Relief from drought and analysis
- Management of livestock.

### ***3.1 Successful ICT Initiatives in Public Sector***

Farmers have always been the back-bone of our country and the Government is striving to strengthen this back-bone of the country through innovative and solid measures. Following are the ICT initiatives in public domain sector.

**AGRISNET:** A foundation network existing at block level working with farming workplaces, horticultural augmentation administrations, and agribusiness exercises to upgrade rustic turn of events.

**Digital green:** The agri. data of neighborhood importance is scattered through advanced video. The framework comprises of a computerized video information base arranged for framers by farmers with the assistance of specialists. The chronicles are appeared to people or little gatherings utilizing workstations, DVD player, TV what's more, to networks through town link organization [13].

**IKSL:** The pertinent data is conveyed to the ranchers on cell phones through five voice message in neighborhood language. Modified arrangements are given to the ranchers through helpline. The farmers can likewise address the specialists on explicit subject through unique "telephone in" programs [14].

**Agmarknet:** This drive gives day by day market cost and appearance data in regard of 300 wares and 2000 assortments in eight nearby dialects. The wide scope of data on costs, appearance, and other related perspectives like evaluations, principles, bundling, and so forth is gathered and dispersed by systems administration major agrarian produce markets working in the country [13, 15].

**iKisan:** is a one-stop answer farmers in giving data on crops, crop the executives strategies, manures, pesticides, and other related data like market updates and climate figures [16].

## **4 ICT Drives Under Non-Governmental Organization (NGO)**

Non-profitable organization assumes a pivotal part in agribusiness and cultivating improvement, generally in provincial and oppressed regions. Progress of checking and appointment among volunteers is the one of a kind key in associate dabs. NGOs carry out constant model and put forth attempt to execute the projects effectively.

Information and Communication tools like e-installment entryways, environment tough models, E-learning, SMS, Denied regions of Vidharba, and Maharastra are major satisfied by NGOs [17]. The important features of these ICT driven support given by the NGOs are.

- Use e-learning tool where farmers are educated to new preservation and ecofriendly items.
- Acts as facilitator between specialist co-op and supplier through web portals.
- Information on digital green from DVDs, videos
- Enhances the result of farmers through web-based media like Facebook, YouTube, and so on
- Complete the gap through mobile calls and SMS
- Solves the questions of farmers identified with polices laid by the public authority.

#### ***4.1 Successful ICT Initiatives in Non-Governmental Organization (NGO)***

**M S Swaminathan Research Foundation (MSSRF):** The foundation expects to speed up utilization of current science for maintainable farming and provincial turn of events. MSSRF centers on ancestral and provincial networks Iin security and climate data. It is a continuous market rate gateway with a single tick local language [18].

**FFMA:** Fisher Friend is a BREW-put together application offered with respect to a minimal expense CDMA handset with a realistic between face, a symbol-based menu and programmable easy route keys. Notwithstanding security and climate data, fishermen can get the areas of fishing regions and constant market costs with a single tick in their neighborhood language [19].

The data important for fisherman is given in neighborhood language through cell phones. The data covered are wave tallness, wind speed and chief, potential fishing zones, relevant news, government plans, and market cost.

## **5 ICT Drives Under Private Sector**

Private sector is tending all the more rapidly to embrace ICT particularly thinking about serious pressing factor in the agriculture sectors. The private sector assumes conclusive parts in India's rural change today, cultivating efficiency upgrades and making occupations and worth in supply chains "from farmers to fork." innovation by the private sector in Indian agriculture as follows.

- Seed/biotech advancements have prompted reported expansions in yields in key field crops, vegetables, and organic products.

- Proprietary crossovers of pearl millet, sorghum, and maize lifted the efficiency of these yields in
- Semi-parched settings not all around served by the Green Revolution.
- Proprietary crossovers cover in any event 75% of the space planted to improved assortments and crossovers.
- Farmers caught generous financial additions from yield expansions in these settings and harvests.
- Private examination has assisted India with expanding fares of harvests, innovation, and agrarian sources of info for example, agrochemicals, and machinery.

### ***5.1 Successful ICT Initiatives in Private Sector***

**Tata KisanSansar:** Consultancy on compost through Tata Kisan Kendra, a call where agro inputs needs for farmers are proposed. IT empower e-visiting focuses are made accessible through video conferencing where specialists examination the sources of info [20].

**Pepsico India:** The Indian outlet of Pepsico makes elegant e-agrarian practices accessible to farmers so as they can raise ranch productivity. 24/7 online interface for questions, call focus kendras are made accessible. GPS empowered answers for topographies and area-based homestead lands. Significant maker of potato like joint effort diverse e-agribusiness like DuPoint, Bayer, and BASF [21].

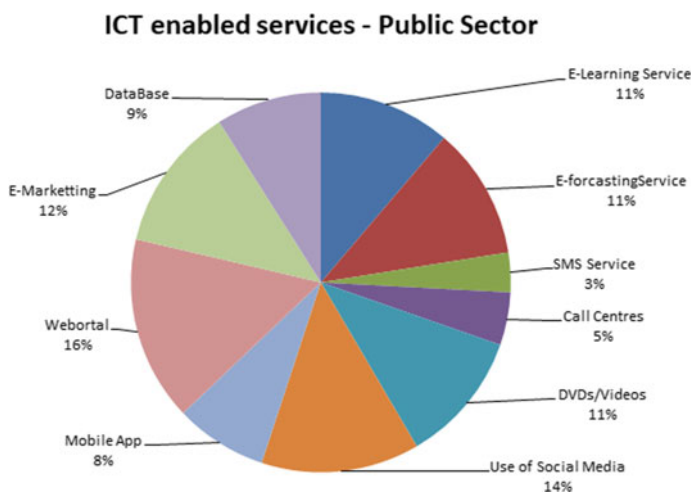
**Mahindara Kisan Mitra:** The start give data on day by day market costs, climate refreshes, crop advisories, agri related news, and so on The data is additionally accessible on different areas like credits, protection, Mandi information base, cold stockpiling and stockrooms, and so on The farmers can likewise get persuaded and take profits by the examples of overcoming adversity of other individual ranchers provided details regarding the site [22].

**E-choupal:** A drive by ITC gives elective promoting channel, data on climate, horticultural practices, input deals, and so forth. It is a stand situated in a town and furnished with PC with web access managed via prepared sanchalak [16].

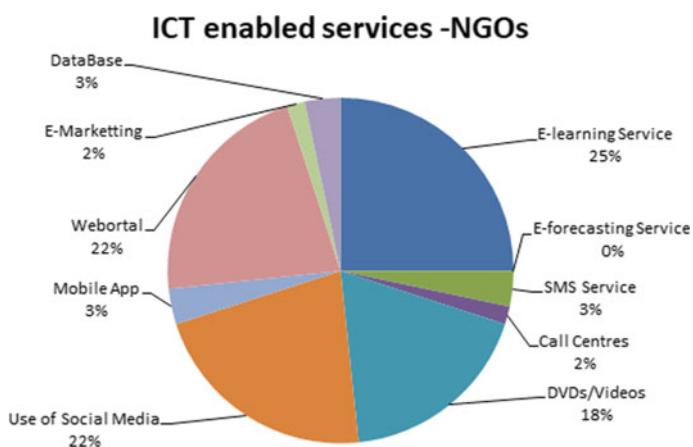
**Warana:** The venture gives admittance to a wide scope of data remembering farming to the individual from the agreeable for nearby language. It gives data on crops, market costs, business plans and instructive freedoms, and so on. The data is given through the town data booths. The administrators of these stands are the principle linkage between the ranchers and the data community [23].

## 6 Comparative Analysis of ICT Services with Respect to Public, NGOs and Private Sector

A summary of ten ICT-enabled agricultural services with respect to Public, NGOs, private sectors surveyed is shown in Figs. 1, 2, and 3, respectively. The greatest percentage of services was for e-learning services, DVDs/videos, and use of social media in all sectors. Services such as e-forecasting services, e-marketing were lower in percentage of services in all sectors. Services such as database, web portal, mobile

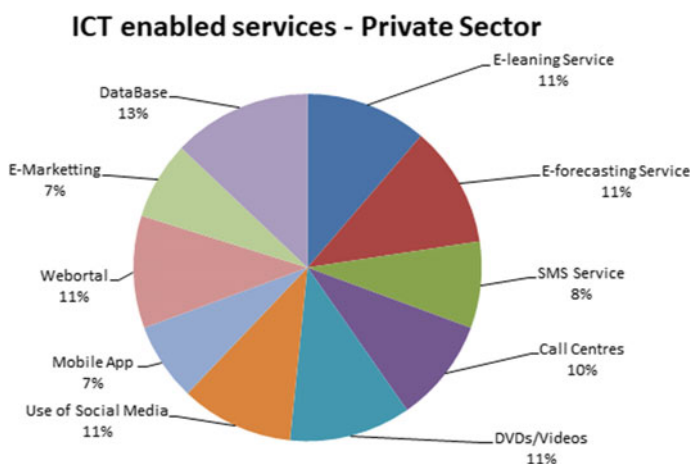


**Fig. 1** Analysis of ICT services with respect to public sector [24–29]



**Fig. 2** Analysis of ICT services with respect to NGOs [24–29]





**Fig. 3** Analysis of ICT services with respect to private sector [24–29]

app, use of social media, call centers, and SMS services were used with moderate percentage.

## 7 Conclusion

India has at any rate two long term insight of utilizing ICT in agribusiness. ICTs have gone through various stages and a significant number of these will keep on developing in light of changing innovation and business climate in agriculture just as because of arising difficulties in agribusiness. Parcel of development has occurred in this area, the expenses of getting to data through ICTs have descended, access have commonly improved and numerous drives are consolidating text, pictures, sound, and video. There is an expanding shift from independent ICT drives to significantly more coordinated ICT drive that essentially works across the worth chain. ICT can assume basic part in reinforcing the limits of farmers as well as for the field level functionaries and middle people. Fostering the privilege or significant substance at the fitting level has consistently been a test, and more endeavors are required toward this path. Strategy pertinent examination is expected to overcome this issue among training and strategy, remembering assessment of effect of ICTs for agribusiness. As there is no stage to share this information and examination on a persistent premise, these are not accessible for investigators and strategy creators to settle on educated choices. Subsequently, there is a need to plan a stage where every one of the major parts in the ICT in farming area will share and utilize information and experience, so that individuals can all in all learn while the drives are in advance and constantly advance.

## References

1. Bahl M (2008) S&T for rural India and inclusive growth: ICT in agricultural marketing. <http://www.nistads.res.in>
2. Nagesh NS, Saravanan R (2019) Impact of ICTs on agriculture growth and development case studies from Karnataka Region. Discussion paper 9. MANAGE-centre for agricultural extension innovations, reforms and agripreneurship, national institute for agricultural extension management, pp 4–5
3. Paul JG, Palaog TD (2019) SMS-based ICT tool for knowledge sharing in agriculture. *Int J Adv Sci Eng Inf Technol* 342–33
4. Reddy, KP, Ankaiah R (2005) A framework of information technology—based agriculture information dissemination system to improve crop productivity. *Curr Sci* 88(12): 1905–1913
5. Sharma JPVS, Roy BR, Dubey SK, Jian MS (2017) ICT based agriculture extension initiatives in India, pp 51–52.
6. Nagendra V (2015) Private sector initiatives in agricultural marketing in India. M.S. Ramaiah University of Applied Sciences Publication
7. Shalendra, Gummagolmath KC, Sharma P (2011) Role of ICT in dissemination of knowledge in agriculture. *Indian J Agricult Econ* 3–5
8. Armstrong J, Gandhi N, Lanjekar K (2012) Use of information and communication technology (ICT) tools by rural farmers in Ratnagiri District of Maharastra, India. In: International conference on communication systems and network technologies, pp 950–955. <https://doi.org/10.1109/CSNT.2012.202>
9. Government of India (2016) State of Indian Agriculture. Ministry of Agriculture and Farmers Welfare, New Delhi
10. Chand R, Singh J (2016) Agricultural marketing and farmer friendly reforms across Indian states and UTs. National Institution for Transforming India. NITI Aayog. pp 2–3; Ferroni M, Yuan Z (2017) Private sector and India's agriculture transformation. *Glob J Emerg Mark Econ* 9(1):1–10
11. Meera SN, Jhamtani S, Rao DUM (2004) Information and communication technology in agricultural development: a comparative analysis of three projects from India. In: *Agricultural research and extension network*, p 135
12. Syiem R, Saravanna R (2015) Access and usage of ICTs for agriculture and rural development by the tribal farmers in Meghalaya State of North-East India. *J Agricult Inf* 6(3):24
13. <https://www.ikisan.com/>
14. <https://www.mssrf.org/>
15. Letshela Z (1999) Developing a web-based agricultural community information centre for rural farmers. *South Asian J Inf Manage* 2(3):4–5
16. <https://www.itcportal.com/businesses/agri-business/e-choupal.aspx>
17. Gangopadhyay P, Khatri-Chhetri A, Shrisath PB, Aggarwal PK (2019) Spatial targeting of ICT-based weather and agro-advisory services for climate risk management in agriculture. *Climatic Change* 154(1):241–256
18. <https://www.qualcomm.com/media/documents/files/india-fisher-friend.pdf-ffma>
19. <https://www.tnagrisnet.tn.gov.in/>
20. Kumar RR (2018) Web based information services for Indian agriculture: a study. *J Libr Inf Commun Technol* 6(2):4–6
21. <https://www.pepsicoindia.co.in/live/story/partnership-with-farmers>
22. <https://www.tatachemicals.com/news-room/press-release/Tata-Chemicals-launches-Tata-Kisan-Sansar>
23. <https://agmarknet.gov.in/>
24. <http://kisanadost.com/KisanMitraWeb/secureuser/login.jsp#b>
25. <http://waranapower.com/sheti-purak-sanstha.html>
26. <https://www.digitalgreen.org/>
27. <https://www.futuregroup.in/>

28. <https://www.indiaseeds.com/>
29. Lathiya A, Rathod A, Choudhary K (2015) Role of social media in agriculture. *Int J Commun Bus Manag* 8(2):268–273
30. Anand S, Satya P, Yedida S, Singh AK (2019) Constraints faced by farmers in access and use of information and communication technologies (ICTs) in Bihar. In: ISEE national seminar-2019 on social digital approaches for transforming Indian agriculture, pp 66–67

# Implementation of e-Birth Registration Systems: Potential and Challenges



## The Case Study of Iran

Mahtab Shahin , Fatemeh Eskandari , Rozha Kamal Ahmed ,  
and Dirk Draheim

**Abstract** This study investigates the feasibility of implementing an e-birth registration system in Iran and identifies existing methods in developed countries such as Estonia. It also identifies how replacing existing methods with digital technology could benefit citizens. In addition, research was conducted to determine what primary requirements are needed to provide e-birth registration services in the Iranian public sector. In this research, the case study strategy employed a mixed-method approach; data has collected and analyzed via triangulation from different sources, i.e., expert interviews, surveys, and document analyses. These findings can provide useful information on different aspects of delivering e-governmental services that need to be enhanced.

**Keywords** e-Birth · Information and communications technology · ICT · Information technology · Unique health identifier · e-Government

## 1 Introduction

According to the World Health Organization (WHO) [16], “the registration of birth is a basic human right, and the birth certification should be conducted immediately after a child is born. In all nations, a birth certificate is a vital legal document that gives an identity to a child” [13]. Furthermore, the information collected from birth registration documents can enable authorities to determine when, how, and which

---

M. Shahin (✉) · F. Eskandari · R. K. Ahmed · D. Draheim  
Information Systems Group, Tallinn University of Technology, Akademia tee 15a, Tallinn, Estonia  
e-mail: [mahtab.shahin@taltech.ee](mailto:mahtab.shahin@taltech.ee)

F. Eskandari  
e-mail: [faeska@taltech.ee](mailto:faeska@taltech.ee)

R. K. Ahmed  
e-mail: [rozha.ahmed@taltech.ee](mailto:rozha.ahmed@taltech.ee)

D. Draheim  
e-mail: [dirk.draheim@taltech.ee](mailto:dirk.draheim@taltech.ee)

sectors to emphasize for development activities [25]. “Civil registration is the way by which countries keep a continuous and complete record of births, deaths and the marital status of their people” [28] and it is employed to issue legal documents, compare also with [19]. Legal identity is among the primary benefits of implementing the integrated identification system, providing citizens with the right to access public services in society. Accordingly, the absence of legal identities under challenging circumstances could adversely affect citizens. Establishing legal identification and registration of infants at birth provides them with specific rights, entitlements, and benefits. Identity documents enable people to claim an inheritance, health benefits, spousal pension rights, and civil rights as land ownership. Proof of identity is essential to nationality requests [1]. In fact, governments provide public service delivery to fulfill the requirements of residents [24]. There is a significant lack of public involvement and participation in the existing model adopted by government agencies to develop and deliver services to citizens. These lacks can lead to various essential issues, such as the implementation of unnecessary or unusable services and distrust in governments due to the lack of reliable processes during service development. In addition, as people’s expectations are rapidly on the rise, citizens expect governments to be more accountable than they currently are and to deliver services efficiently and effectively [5, 7].

Information and communication technology (ICT) [11] enables governments to provide better and more effective services for citizens and enterprises [4, 27]. In fact, capabilities and technological advancements of ICT encourage governments toward reforming into implementing e-government [2, 5]. Before the definition of public e-services, an e-service should first be defined. According to [17], operating in the electronic services of the public sector is defined as a dynamic process aiming at value fulfillment for clients. By definition, an e-service refers to an electronic service, i.e., the services that are delivered electronically. Regarding the concept of e-government, e-services generally deal with intangible products such as the information exchange to obtain approvals, payments, and registration tax. According to some analysts, e-services are the information-based processes of interactions between governments and citizens [17].

Although governments are increasingly using e-services, the extent of implementation varies from country to country [3]. Differences demonstrate that there is slow e-government progress in developing countries, some of which are even experiencing regression due to facing numerous dynamic drawbacks. The reasons could lie in gaps in the Internet technology infrastructure, procedures, utilization, and sufficient resources to construct the costly national communication infrastructure [3]. This is because the developed countries are pioneers of e-services [14]. In recent years, all developed and even developing countries have taken successful steps toward the deployment of e-governments in order to enhance people’s quality of life [9], improve their satisfaction [26], and boost governments and businesses in relation to their national economic dynamism [10, 18]. In addition, by increasing people’s awareness [22] of modern technologies and electronic services, governments are expected to provide public services for various segments of society more efficiently than ever before at the lowest cost and shortest time [2, 5]. Therefore,

these international experiences could be used as helpful resources to provide new ways of delivering services in the Iranian e-government system.

The rest of the paper is organized as follows. In Sect. 2, we provide an overview of Iranian birth registration by details. In Sect. 3, we explain the research methodology. In Sect. 4, the results from the interviews and questionnaires are given plus an analysis of the e-birth registration in Estonia. In Sect. 5, we discuss our results and provide recommendations. In Sect. 6, we discuss related work. Finally, we finish the paper with a conclusion in Sect. 7.

## **2 Overview of Birth Registration in the Case of Iran**

Based on the Iranian Registry Act (Birth Registration in Iran, An analysis of the state of relevant laws in Iran, 2005), when a child is born in a hospital or a maternity ward, parents should take the certificate from the hospital with an ID card or resident permit. The hospital should also send the report to the National Organization for Civil Registration (NOCR) database. In the next step, parents bring the certificate to the civil registration office, fill out the paperwork, and make the required payment to apply for the birth registration under some circumstances stating that parents must be married and that the announcement of birth registration should be done by the father or the father's father. Within 15 d, one of the parents should then visit the office again to collect their child's birth documents. However, the process is a little more time-consuming for those who live in remote/rural areas. In fact, the process starts with the reports of birth by local trustees, district governors, teachers, or post officers in the country where they are working as executive agents of governmental entities in their respective regions to the nearest local/office registration department. They collect the reports and required documents and send them to the governmental birth registration center where, after the identity is confirmed, the birth document is issued and sent back to the local department office in the respective regions. Finally, the birth document will be collected from local trustees, district governors, teachers, or post officers to hand it over to parents who are living in rural areas. The current process needs to be operated in more efficient and effective manner and that can be achieved through digitization of the business process.

## **3 Research Methodology**

This research employs an exploratory case study strategy as it allows for an in-depth investigation of a contemporary within its real-life context [29]. The case study strategy aims at exploring and investigating the topic where no earlier study has been conducted [23, 29]; therefore, it suites the current study to investigate the case of Iran to present status of birth registration in the country and the challenges in implementing an e-birth registration, with exploring the potentials of future implementation of e-birth registration system.

We used a triangulation of multiple sources for data collection to strengthen a validity of the study [29] through a mixed method of qualitative and quantitative approaches. In qualitative approach, we used expert interviews as it is considered a crucial source of data in case study research [23, 29] due to the fact that the knowledge gathered from the people closely involved in the case will provide researchers with much detailed information [23]. The interviews were approx 20min long, and four hospital employees were interviewed to provide their in-depth views on the current status of the Iranian birth registration system at the hospital and their perspective of future digital transformation. We considered an RQDA for analysis of interview responses [8], and used a thematic analysis for the coding process [6] in RQDA. Additionally, the relevant available literature was analyzed to obtain information on the current applications of e-birth registration systems as analysis of the related work to the subject considered to be valuable for the case study research [29].

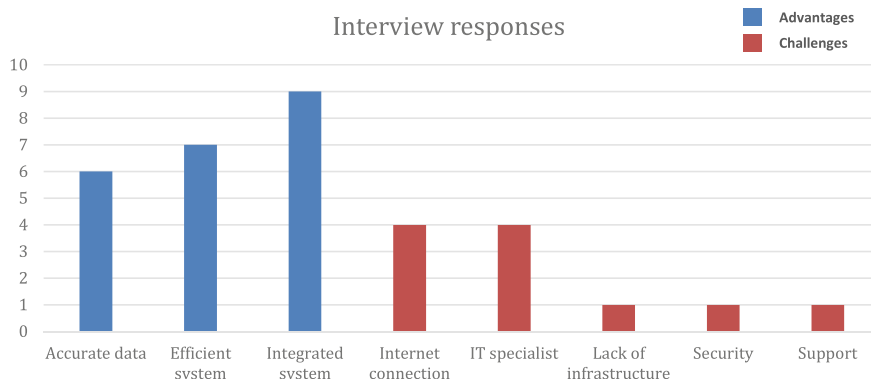
Furthermore, in quantitative approach we distributed online surveys among citizens to get overview on five dimensions, i.e., demographics, levels of awareness of citizens, the interest of citizens in governmental portals, levels of satisfaction of citizens based on changes, and levels of trust through 19 sub-criteria, and 158 responses were collected from surveys.

## 4 Results

### 4.1 Interview Results

Four expert interviews were conducted with hospital employees from two different cities of Iran, namely Zanzan and Esfahan, to determine the problems and barriers between the civil registration center and childs' birth certificate units in addition to providing their views on the future implementation of the electronic birth services. The result of analysis is provided in Fig. 1, divided into two sections as advantages of the implementing e-birth registration system, and challenges of implementing it.

Interviewees started by presenting the problems in the current system and challenges associated with issuing a birth registry certificate in hospital, as the process is performed very slow due to staying in long queues to obtain a birth certificate and gets more complicated when it involves legislation problems with marriage certificates of the parents. Analysis showed that respondents considered implementing e-birth system as a vital integrated system to reduce the current bureaucracy rate and simplify the process, as can be seen from the figure that the code integrated system is been stated by all 4 respondents 9 times. Additionally, respondents 7 times stated that electronic process of birth registration will increase the efficiency in the system and allows all users to perform operations more smoothly. Another advantages of the e-birth system as considered by interviewees were having accurate data regarding birth certificate as it has been stated 6 times by respondents and considered accurate data as a crucial point for satisfying clients and government simultaneously and



**Fig. 1** Analysis of interview responses

ensuring the correctness of the process. Next, interviewees also explored the future challenges that could face the implementation of e-birth system and were mostly focusing on the lack of IT specialists and instability of Internet connection that has been stated equally 4 times by all respondents. Other concerns were relevant to the lack of IT infrastructure in the Iranian government, and need of constant support through providing training sessions, as well as security issues of the system also stated to be maintained at a very high level.

## 4.2 Survey Results

The survey analysis of Iranian citizens illustrated that approximately 90 percent of respondents was aged 20–40 years old, and that they mostly had higher education degrees. Therefore, most of the respondents were at least able to use online services. Although a high usage rate of the Internet was estimated at nearly 97% daily, the significant Internet usage has belonged to social media and Web surfing. However, the public/private online service usage accounted for 10% daily, and only 42% of respondents confirmed that they were using e-services regularly. Moreover, some of the Iranian citizens are not familiar with the advantages of online services, also; on the other hand, answering the sixteenth question, the large number of responses admitted that they preferred to use online services; however, the responses to another question indicated that they usually spent time in queues to access governmental services in administrative offices. The other point is that only less than half of the responses were satisfied with the information accuracy and the use of up-to-date information in public e-services. As a result, this demonstrates that a few citizens would be likely to use electronic services if they were sure about security, ease of use, availability of sufficient knowledge in the use of e-services, and information accuracy. Moreover, many of them commented on the Internet connection speed and



quality of e-service platforms, which means that there is high demand for developing the primary infrastructure in the Iranian e-government system and implementing e-birth registration system to increase efficiency and effectiveness in the process.

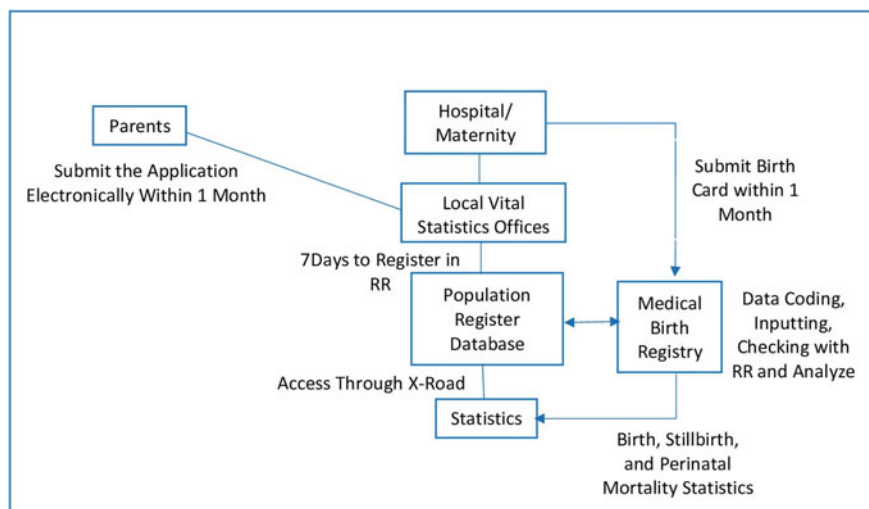
### ***4.3 Comparison with the Case of Estonia***

To implement e-birth registration services effectively in the Iranian civil registration system, it is necessary to utilize the frameworks and backgrounds of other countries which are currently providing such services for their citizens. Hence, Estonia can be considered the best case as a country with a high-ranked index in the e-governance context for implementing electronic services [20] and would like to use technology to help improve life for its citizens [12]. The first instance of identification (i.e., a medical birth certificate) is performed by physicians either at childbirth at the hospital or a prenatal check. This first child record is carried out as a birth certificate, which is the vital document for the Estonian Medical Birth Registry (EMBR) system. A birth card is allocated to every child born in Estonia. The cards are sent to the registration system monthly in a fast way. This information is obtained from all maternal health providers. The birth card is mainly characterized by the personal identification code (PIC), issued to each newborn at the time of birth based on gender, date of birth, and a three-digit code indicating the order of birth in the total number of daily deliveries. These codes change daily and are circulated to hospitals in lists. If the child has not been born in the hospital, the parents should ask the healthcare provider to complete the birth registration card, or the vital statistics officer will grant the PIC at the time of birth registration [1] (Fig. 2).

In order to record birth, parents must send an application to the vital statistics office. The officer is responsible for registering the birth within seven working days after the request. Birth registration and the production of a birth certificate are free of charge, and the application is not required to be sent in person. However, the application is submitted digitally and must contain the following pieces of information: name, gender, date of birth, personal identification code, place of birth and citizenship of the child, personal identification code of the mother, personal identification code of the father, right of custody.

Regarding the Estonian system integration, it is essential to mention that the government uses this systemic view to establish all these components to enhance the delivery of public services and maintain the effectiveness of public administration.

As a result, the integration process followed a hierarchy where both components had a purpose and cooperated. In addition to the perceived strategies mentioned, three key important factors were providing the basis for system integration: comprehensive policy and legal frameworks, public key infrastructure (PKI) [4], and public-private partnerships (PPP) [1].

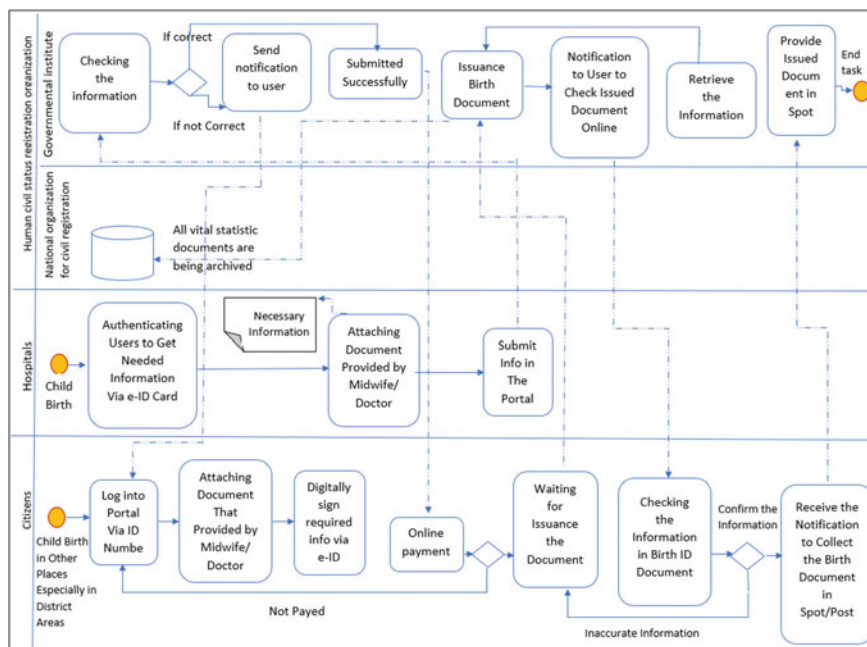


**Fig. 2** Direction of e-birth registration process in Estonia [1]

## 5 Discussion and Recommendations

The conducted survey and interviews helped evaluate citizens' awareness and their tendencies to use governmental e-service and their perspective regarding e-birth registration services. The survey also addressed the obstacles in the current system and the main demands for providing e-services. In addition, studying the Estonian e-birth registration system could be the best practice to provide the necessary frameworks for the Iranian birth register system.

According to the results, the main challenges in the first step included providing the components of the primary infrastructure such as the high-speed Internet connection without disruption, employing IT specialists in hospitals to deal with any technical issues, and also providing the technological infrastructure and the integrated system to enable authorities of hospitals and those of the national birth registration database to exchange information on a most secure and trackable platform. This model no longer needs local departments and offices and local trustees as an intermediary organization, and as a consequence, this could reduce the employee workloads. Another point is the current legislation problems in the current registration process. According to the final question in an interview, it is necessary to revise the type of marriage, parental nationalities, and unmarried parents by considering changes due to their negative effects on the acquisition of birth documents. Iranian respondents were partially aware of the theoretical e-governance process and were eager to use the alternative way of enhancing and easing their lives. However, some people, especially in rural areas, lack the necessary knowledge about e-services or have no access to the Internet connection. Therefore, meaningful steps should be taken to provide an efficient and effective way of delivering electronic services by



**Fig. 3** To-be model for Iranian birth registration system

authorities. Another important point that needs to be discussed is identification with e-ID cards. As mentioned by interviewed employees, providing an integrated portal with a civil registration organization could simplify the process, so e-ID and ID card readers are needed to facilitate this process and enable citizens to sign digital documents as a public e-service implemented in Estonia. Although most Iranian citizens have received their e-ID cards, due to reasons (mismanagement, lack of primary infrastructure, US sanction against Iran's government), this project is, in reality, after a decade, failed to achieve the planned goal. Overall, these findings show how to make the most of e-service provisions in simplifying implemented governmental services.

The to-be model, shown in Fig. 3, has been designed to help visualize the implementation of the e-birth registration system in Iran. It has been developed a modernized and simplistic way for citizens to access birth registration service by considering how this service has been implemented in developed countries such as Estonia. Such a system could simplify the issuance of the birth certificate process, reduce the costs and time consumption for citizens, particularly those living in remote areas, and mitigate the administrative burden. In addition, this e-service can obtain accurate statistics of the birth rate in real time and improve in providing comprehensive information to other sectors.

The event in this model starts from the hospital. After the newborn report, by utilizing e-ID cards, users would be authenticated without providing previously needed

documents (marriage document, ID-document booklets). Further, documents that the doctor/midwife provides would be attached and submitted to the Iranian Civil Registration Organization. Furthermore, it should mention that those living in a remote district or their birth child do not take place in the hospital. The starting event is when the parent logs into the portal via ID code, attaching a needed document, and signs documents digitally. After submitting information in the portal and confirming that documentation is submitted successfully, the user should process the payment successfully and wait for the issuance birth document. Eventually, the citizen gets a notification to check and confirm issued documents in the portal. Then, they could collect the birth document on the spot or via postal service, and the events end.

## 6 Related Work

There are several methods that countries use to recognize individuals in order to allocate health care [30], which is as follows:

- In some countries dedicate both national unique health identifier (UHI) and unique identification number (UIN) to each individual.
- in some other countries, the UIN as a unique number is used also for health purposes.
- In other countries, individuals are assigned a UHI without having a UIN.

Among the high-income countries, Slovenia [30] as a European union uses both UIN and UHI. These two numbers are linked in the central population register (CPR). In this country, birth registration starts when a child is born. On the other hand, before the discharge of the family, the birth registration data sends to CPR, which then creates a UIN for newborns. CPR shared the data with the Health Insurance Institute, and UHI and tax numbers are created. In addition, Thailand as an upper-middle income country [21] and South Korea as high-income countries are used UIN as UHI.

In 1982, the civil registration system in Thailand moved to digital work from paperwork, which was simplified by assigning a UIN to each Thai citizen, known as a personal identification number (PID), and then entering the citizen information into the general database population. The hospital presents a birth notification form to the newborn's parents, and then they take the notification form to the district registration office along with his/her proof of identity to register the birth of the child. The registering clerk can certify the submitted information to register the child, issue the birth registration, and allocate a PID. Since then, the PID is used for identifying the Thai child and accordingly enrolled into the free national health insurance [21].

In 1960, the CRVS system was developed in South Korea. This system was consist of three main parts: family relationship registration (FRR), resident registration (RR), and vital statistics. The system generates an RR number for the RR system during the birth registration as a section of FRR. The RR card is issued for all 17 years and older citizens. This card individual's number is used for the connection of different databases, for instance, income, tax, property, and family relationships [15].

## 7 Conclusion

In general, the adoption of e-government can result in several aspects of governing between the government and citizens as well as the government and the private sector by decentralizing governmental data, increasing effectiveness and productivity, enhancing information transparency, improving citizen satisfaction, and reducing the complexity of service delivery and bureaucracy. However, despite all the efforts and strategies to enhance e-government in Iran, achieving e-governance objectives has been complicated and is associated with several challenges.

This paper aimed to determine how the Iranian government and citizens could benefit from implementing e-birth registration services. According to the research findings, replacing the traditional services with e-services increased public and citizen satisfaction by delivering high-quality public services and reducing the complicated bureaucratic process. In fact, the current birth registration system is performing manually by citizens and several numbers employees in different organizations to accomplish the issuance birth documents. So it can be seen that by transforming this service to an online way by providing a civil registration platform, citizens will submit and obtain needed documents in a fast and user-friendly, and cost-efficient way. Furthermore, authorities would be ensured that data could be accessed, collected, and stored in a real-time framework. Also, based on the interview and the survey in this paper, the respondents emphasized that implementing e-services saved time and resources, removed long queues, and eliminated face-to-face visits.

Besides that, challenges that governments and citizens might face while utilizing this service are the considerable costs of implementing e-governments. The other point that responders in the survey emphasized was the weaknesses of the required infrastructure, such as Internet access, speed, and coverage across the country. Also, regard to the level of citizen's awareness as one of the essential factors in implementing e-services which could impact on principles of e-participations. Although the survey results showed that public awareness of e-governance utilization is relatively high, their responses in open-end questions indicated that the authorities need to improve the level of awareness among citizens more than before.

**Acknowledgements** This work has been partially conducted in the project "ICT programme" which was supported by the European Union through the European Social Fund.

## References

1. Aguilar Rivera AM, Vassil K (2015) Estonia: a successfully integrated population-registration and identity management system. World Bank
2. Ahmed RK, Lips S, Draheim D (2020) eSignature in eCourt systems. In: Proceedings of 2020 fourth world conference on smart trends in systems, security and sustainability (WorldS4). IEEE, pp 352–356

3. Ahmed RK, Muhammed KH, Pappel I, Draheim D (2020) Challenges in the digital transformation of courts: a case study from the Kurdistan Region of Iraq. In: Proceedings of ICEDEG 2020—the 7th International Conference on eDemocracy & eGovernment. IEEE, pp 74–79
4. Ahmed RK, Muhammed KH, Reitsakas A, Pappel I, Draheim D (2019) Improving court efficiency through ICT integration: identifying essential areas of improvement. In: Proceedings of ICT4SD 2019—the 4th international conference on ICT for sustainable development. Lecture Notes in Networks and Systems, vol 93. Springer, pp 449–461
5. Ahmed RK, Muhammed KH, Pappel I, Draheim D (2021) Impact of e-court systems implementation: a case study. *Transf Gover People, Process and Policy* **15**(1)
6. Braun V, Clarke V Using thematic analysis in psychology. *Qual Res Psychol* 3:77–101
7. Butt SA, Pappel I, Oolu K (2021) Implementation of electronic records management systems: potential and challenges a case study of the water and power development authority (wapda) in pakistan. In: *ICT analysis and applications*. Springer, pp 629–639
8. Chandra Y, Shang L (2017) An RQDA-based constructivist methodology for qualitative research. *Qual Market Res: Int J* 20(1):90–112
9. Draheim D Blockchains from an e-governance perspective: potential and challenges
10. Draheim D (2020) On architecture of e-government ecosystems: from e-services to e-participation:[iiwas' 2020 keynote]. In: Proceedings of the 22nd international conference on information integration and web-based applications and services, pp 3–10
11. Draheim D, Krimmer R, Tammet T (2021) On state-level architecture of digital government ecosystems: from ict-driven to data-centric. In: *Transactions on large-scale data-and knowledge-centered systems XLVIII*. Springer, pp 165–195
12. Dreyling R, Jackson E, Tammet T, Labanava A, Pappel I (2021) Social, legal, and technical considerations for machine learning and artificial intelligence systems in government
13. Gerber P, Gargett A, Castan M (2011) Does the right to birth registration include a right to a birth certificate? *Netherlands Q Human Rights* 29(4):434–459
14. Hassan H, Shehab E, Peppard J (2011) Recent advances in e-service in the public sector: state-of-the-art and future trends. *Bus Proc Manage J*
15. Kang M, Bae G, Kim H, Hong SY, Mills S (2019) Korean resident Registration system for universal health coverage. World Bank
16. Lee K (2008) *The world health organization (WHO)*. Routledge
17. Lindgren I, Jansson G (2013) Electronic services in the public sector: a conceptual framework. *Government Inf Q* 30(2):163–172
18. Lips S, Bharosa N, Draheim D (2020) Eidas implementation challenges: the case of estonia and the netherlands. In: *International conference on electronic governance and open society: challenges in Eurasia*, pp 75–89. Springer
19. Mahapatra P, Shibuya K, Lopez AD, Coullare F, Notzon FC, Rao C, Szreter S et al (2007) Civil registration systems and vital statistics: successes and missed opportunities. *Lancet* 370(9599):1653–1663
20. McBride K, Misnikov Y, Draheim D (2021) Discussing the foundations for interpretivist digital government research
21. Mills S, Lee JK, Rassekh BM, Kodolja MZ, Bae G, Kang M, Pannarunothai S, Kijsanayotin B (2019) Unique health identifiers for universal health coverage. *J Health Popul Nutr* 38(1):1–8
22. Pappel I, Butt S, Pappel I, Draheim D (2021) On the specific role of electronic document and record management systems in enterprise integration. In: *Proceedings of fifth international congress on information and communication technology*, pp 37–51. Springer
23. Runeson P, Host M, Rainer A, Regnell B (2012) *Case study research in software engineering guidelines and examples*. WILEY
24. Simeonov PL, Ehresmann AC, Smith LS, Ramirez JG, Repa V (2010) A new biology: a modern perspective on the challenge of closing the gap between the islands of knowledge. In: *European conference on a service-based internet*. Springer, pp 188–195
25. SlothNielsen J, SlothNielsen R (2020) Mothers and others: Transgender birth, birth registration and the rights of the child, with a focus on the united kingdom and south africa. *Int J Discrimination Law* 20(4):203–223

26. Tsap V, Lips S, Draheim D (2020) Analyzing eid public acceptance and user preferences for current authentication options in estonia. In: International conference on electronic government and the information systems perspective. Springer, pp 159–173
27. Verdegem P, Verleye G (2009) User-centered e-government in practice: a comprehensive model for measuring user satisfaction. *Gov Inf Q* 26(3):487–497
28. World Health Organization: Civil registration: why counting births and deaths is important (2014). <https://www.who.int/news-room/fact-sheets/detail/civil-registration-why-counting-births-and-deaths-is-important>
29. Yin RK (2014) Case study research: design and methods. SAGE, 6 edn
30. Zorko Kodolja M, Mills S (2019) Slovenian civil registration and unique identification number system for universal health coverage: a case study

# Risks Reduction of Rainfall-Induced Landslides-A Site-Specific Early Warning System (SSEWS)



Martin Kuradusenge , Santhi Kumaran, Marco Zennaro ,  
and Jean Baptiste Minani 

**Abstract** This paper describes a Site-Specific Early Warning System (SSEWS) for rainfall-induced landslides using the soil moisture sensors deployed in the regions prone to landslides' incidences to minimize the risks caused by this natural threat. An experimental study was conducted to estimate the soil moisture level that induces a slope failure. The estimated thresholds were used in the development of the SSEWS prototype. The system on site consists of a sensor node that gathers the soil moisture data from the remote sites (landslide prone areas). The collected data are transmitted using a GSM/GPRS (Global Systems for Mobile/General Packet Radio Service) module over the cellular network to the database server and analyzed for issuing alert and visualized on a dashboard. The system was tested on different sites and shown to be 71.4% successful.

**Keywords** Rainfall · Soil moisture content · Sensors · LEWS · Slope failure

## 1 Introduction

Rainfall-induced landslides has been challenging the world due to the loss they cause on lives, natural environment and infrastructure. Universally, different strategies and techniques have been used to mitigate or reduce the risks caused by this natural threat.

---

M. Kuradusenge (✉)

African Centre of Excellence in Internet of Things, University of Rwanda, KN 67, Kigali 3900, Rwanda

e-mail: [m.kuradusenge@ur.ac.rw](mailto:m.kuradusenge@ur.ac.rw)

S. Kumaran

School of ICT, The Copperbelt University, 21692 Kitwe, Zambia

e-mail: [santhi.kr@cbu.ac.zm](mailto:santhi.kr@cbu.ac.zm)

M. Zennaro

International Centre of Theoretical Physics, Strada Costiera, 11, 34151 Trieste, Italy

e-mail: [mzennaro@ictp.it](mailto:mzennaro@ictp.it)

J. B. Minani

Concordia University, 12157 Guy St, Montreal, QC, Canada



Landslide mitigation strategies can be categorized into structural and nonstructural strategies. Structural strategies are like drainage, walls, planting, directing, etc. Their implementation may be challenging in developing countries due to the economic constraints. Non-structural measures are consequently reducing measures, which may include land use planning, public willingness, emergency management, early warning, etc. [1]. Digital technologies appear to be the most effective landslide risks mitigation methods used in different countries. The use of the digital technologies for landslide risk reduction comprises of digital mapping, landslides prediction and early warning systems [1]. One of the digital technologies used nowadays is the Internet of Things (IoT). IoT supports quick data collection from the remote site, transmission of the data, analyze, and feed them to the early warning system for alerting the people under risks. L. Piciullo et al. have discussed a landslide early warning systems (LEWS) which has proved efficient in providing information about the imminent disaster, so that precautions can be taken before the incidence occur and hence reduce the risks that would cause [2].

Globally, the schematic representation of early warning system proposed by the United Nations Office for Disaster Risk Reduction (UNDRR) and World Meteorological Organization (WMO) comprises of four components: (i) disaster risk information, (ii) monitoring and analysis of threats, (iii) communication and warnings, (iv) readiness, and response abilities [3, 4]. Generally, monitoring and warning can be considered as the main components of LEWS [5].

- The monitoring process involves all operations related to the in situ installation of the instrumentation, data acquisition, processing, and analysis [6]. In situ installation involves on-field placement of measuring instruments such as meteorological stations, GPS systems, webcam, sensor nodes (inclinometer, piezometer, soil moisture, etc.) [7] to gather targeted data to be later processed for LEWS.
- The data acquisition procedures involve setting the measurement frequency (cycles). The data are periodically captured by the monitoring instrument according to the preset cycle and transform them in the digital format. This process is performed locally and is terminated by the local export, which involves the transmission of data from the measuring node to the nearest gateway, and the central database (server).
- The data transmission involves short range such as Bluetooth, Low Power Personal Area Network (6LoWPAN), etc. [7]; the long range data transmission uses technologies like LoRa (Long Range), SigFox, and cellular network technologies such as GPRS, 3G, LTE (Long Term Evolution) [8].
- The warning procedure involves the process of comparing the actual data to the preset threshold derived from the historical data collected through various experimentation.

## 2 Related Works

Organizations in different countries have implemented LEWS using different techniques, whereas the research and studies are still ongoing to find out a reliable solution for notification of landslide incidence prior to the occurrence. LEWS can be categorized as global, territorial or local systems, operating at wide, regional, or local area scale, respectively [1]. The regional or territorial LEWS predicts the rainfall-induced landslide occurrence over a large region by monitoring meteorological parameters and providing an alert to the population and authorities in charge of disaster management [2, 9]. Examples of regional LEWS are the Hong Kong Landslip Warning System, LEWS for the San Francisco Bay area (USA), and LEWS for Java (Indonesia) as described in [10]. The local LEWS monitors parameters inducing the landslide occurrence at a small well-identified slope and provide warnings to the stated people [1]. Åknes rock slope in Norway is a typical case of this category [11]. The regional LEWS utilizes weather stations, radars, and other technologies that are expensive so that non-developed countries cannot afford. Besides, the coverage areas of LEWS can be a challenge of accuracy and reliability. Local LEWS can overcome different challenges of territorial LEWS.

The reliability of LEWS depends on a good rainfall threshold, which is appropriate for territorial LEWS (Te-LEWS) [9]. However, the prediction capability is limited because of dependence on the threshold identified from historical data. The main problem of the last is the accuracy due to the poor spatial distribution of rain gauge stations. Therefore, the local LEWS requires the site-specific rainfall data.

The purpose of this study is to design and develop a tailored cost effective prototype for Site-specific LEWS (SSLEWS). We used the soil moisture data, as recent studies showed that the hydrological threshold can improve the prediction capability [10]. Thresholds for rainfall and soil moisture have been determined through field experiments that found threshold values differ from one site to another depending on the geo-environmental characteristics such as slope, soil types, and land cover.

## 3 Data and System Architecture

### 3.1 *Rainfall and Soil Moisture*

Regional LEWS considers different rainfall characteristics such as current and antecedent rainfall, intensity, and duration. The local LEWS monitors landslide events using some more parameters based on personal experience (or expert judgment) [11]. Soil moisture has been proved to be one of the most useful parameters [11–16].

To overcome this problem, we conducted an experimental study to determine rainfall thresholds for different areas. Different sites have been chosen according to the various geo-factors including the slope inclination, soil types, and land cover/use.

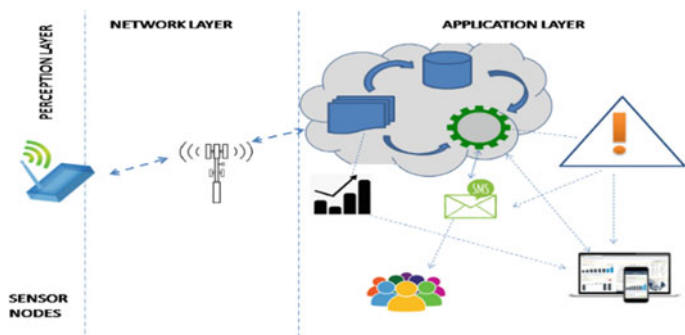
We used the rainfall simulator to provoke slope failure, and rain gauge was used to capture the rainfall amount in time series.

Although the hydrological thresholds are not much used in designing the LEWS, recent studies showed that the in situ measurement of soil moisture have a crucial importance on improvement of LEWS [10, 12] because this parameter has specific information on local landslide activity [14]. The soil moisture thresholds on specific sites were identified concurrently with that of rainfall as explained in the previous paragraph. Three sensors were placed underground in various depth (20, 70, and 120 cm). All data collected during field experiments are available online [17]

### 3.2 System Design

The site-specific landslide early warning system (SSLEWS) is made through a wireless sensor network (Internet of Things) that uses sensor nodes, which are spatially distributed to collect soil moisture data, transmit them to the nearest base station of the cellular network, then to the cloud for landslide prediction process, storage, visualization, analysis, and alert. The system is made of three main layers: perception, network, and application as shown by Fig. 1.

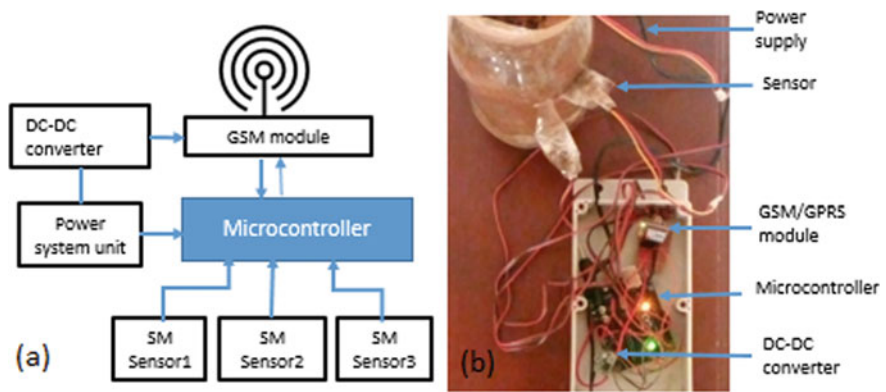
**Sensor node.** Each sensor node comprises the power supply, three soil moisture sensors, microcontroller, and GSM/GPRS module. The power supply unit consists of off-grid system comprising solar panel, charge controller, battery, inverter, and enclosing accessories. The sensing part involves the capacitive soil moisture sensor v1.2 for measuring the ground water content. The sensors were calibrated to give the sensed data in the range of 0–100. The soil moisture sensors were interfaced with microcontroller (Arduino UNO) for central processing. The GSM/GPRS (SIM800L) module was used to connect the sensor node to the server via cellular network. The



**Fig. 1** Proposed SSLEWS network architecture. The perception layer consists of sensor nodes. Network layer consists of a gateway (base station) and other interconnecting devices of cellular network. Application layer comprises data acquisition, storage, visualization, analysis, and alert

power unit and GSM module are interfaced by DC to DC converter for adapting 12 V from power unit to 4 V needed for GSM/GPRS module (Fig. 2).

**Application.** The application layer consists of data storage, processing, analysis, visualization, and alert. The new data received from network layer are stored in the database where they can also be visualized in real time. The incoming data are then compared to the threshold to classify the warning level (Table 1).



**Fig. 2** a Block diagram of the sensor node, b inside the box and calibration

**Table 1** SSLEWS warning levels

Warning level	Description	Action
Level 1: Normal	No threshold attained	No action should be taken. Authorities/responsible agents can visualize the data
Level 2: caution	One of the thresholds is attained. There is a possibility of landslide occurrence	Alarm should be sent through short messages on mobile phones of local citizens. Data and warning can be visualized by authorities or agents in charge
Level 3: Critical	Both thresholds are attained. There is a high possibility of landslide occurrence	Local siren should be activated and alarm message is sent through short messages on mobile phones of local citizens. Data and warning can be visualized. People should be evacuated

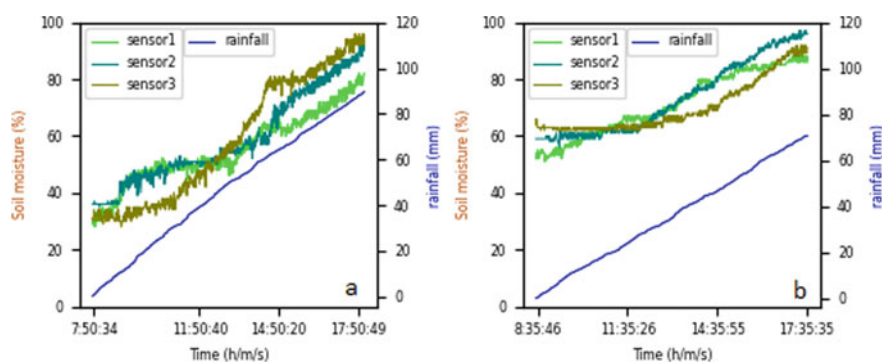
## 4 Experimental Results and Discussion

The prototype setup was preceded by an experimental study to identify the correlation between rainfall and soil moisture content. The study was conducted at various sites characterized by different factors such as soil types, slope, and land cover. This activity carried out by using a rainfall simulator on different sites while recording the rainfall amount by rain gauge and soil moisture by capacitive soil moisture sensors. It took about 8–11 h with a rainfall intensity of 7–11 mm per hour to induce landslides on the selected sites. The result shows that there is a linear regression between rainfall and soil moisture content as shown in Fig. 3.

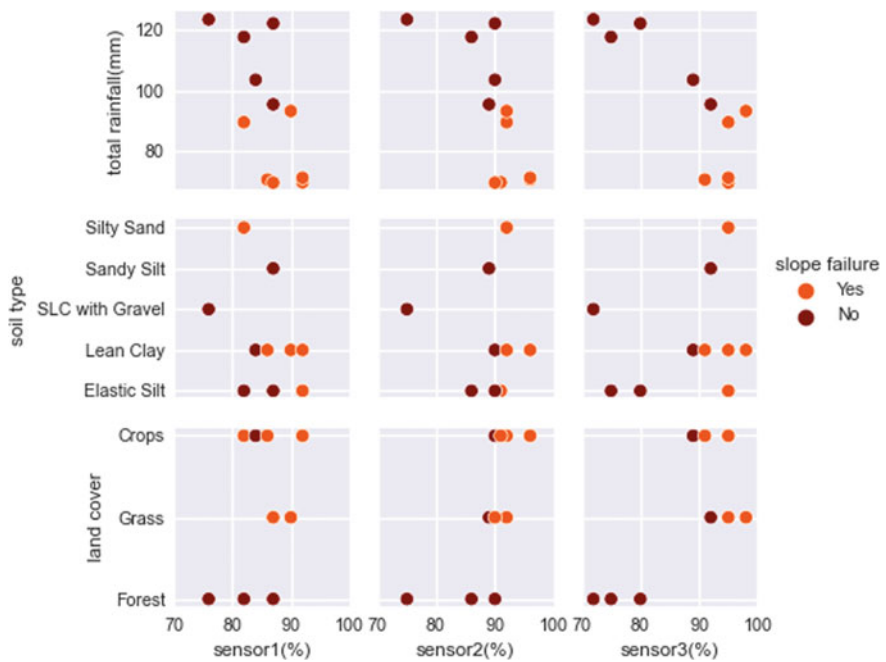
Experiment study was conducted at 11 sites. Depending on the site specifications, some sites indicated the sign of slope failure, others did not. Six sites shown the sign of land sliding. Slope failures that occurred at different levels of rainfall and soil moisture are as shown by Fig. 4.

The estimation of rainfall and soil moisture inducing landslides could be drawn from the plots in (Fig. 4). The experiments revealed that slope failure occurred when the soil moisture content was at least 90% for the two sensors placed between 50 and 120 cm depth, whereas the top sensor (20 cm) recorded the minimum of 80%. It was also observed that Lean Clay and Elastic Silt are prone to slope failure as well as the land covered by crops or grass. The prototype was tested to the land covered by crops, soil type of Lean Clay, and the following threshold were used to provide various warning levels as described in Tables 1 and 2.

Seven sites were selected to test the prototype. The test was done by wetting the soil above the cut slope above the house until a portion of the wetted soil falls (or indicate the sign falling). Out of seven sites, five slopes shown the sign of failure (71.4%) while two sites could not fall even if the soil water content was above 92% for all sensors. The soil moisture contents recorded at each site are shown in Table 3.



**Fig. 3** Correlation between rainfall and soil moisture. **a** Features characterizing those sites were Silty Sand as soil type, the slope inclination of 26%, land covered by crops. **b** site features were Lean Clay, slope inclination of 29%, land covered by crops



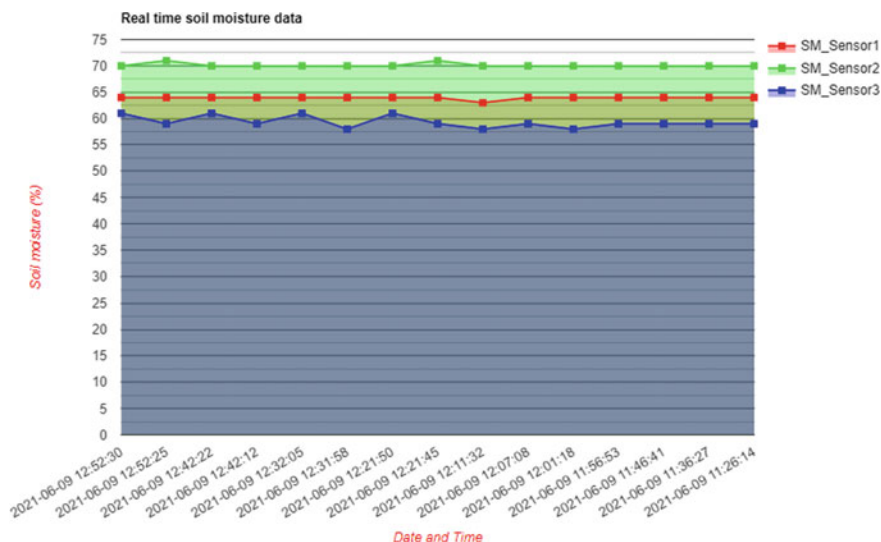
**Fig. 4** Slope failure versus sites’ parameters. The first row plots rainfall versus soil moisture; the second row plots soil types versus soil moisture; the third row plots land cover versus soil moisture

**Table 2** Warning levels and soil moisture thresholds

Warning level	Soil moisture thresholds		
	Sensor1(20 cm)	Sensor2 (70 cm)	Sensor3 (120 cm)
1	≤60	≤70	≤70
2	≥60	≥70	≥70
3	≥80	≥90	≥90

**Table 3** Soil moisture content at slope failure on the tested sites

Sensor	Site1	Site2	Site3	Site4	Site5	Site6	Site7
1 (%)	88	96	93	89	93	95	83
2 (%)	97	99	97	88	96	92	89
3 (%)	94	99	95	93	99	99	92
Slope failure?	Yes	No	Yes	Yes	No	Yes	Yes



**Fig. 5** Real-time soil moisture data from the system dashboard

The prototype of SSEWS was placed above the cut slopes of the homes with historical background of slope failures. The sensor nodes collect the soil moisture contents and transmit them to the remote database. In the real time, data can be visualized and analyzed over the system dashboard as indicated by Fig. 5. In addition, alert can be issued to the people in the area through short message service (SMS) and only warning level 2 and 3 can be received via SMS, whereas level 1 does not need any caution.

Every ten minutes the new data entry from sensor node is recorded and stored into database for later use and analysis. The system dashboard can be visualized online.

## 5 Conclusion

The main objective of this research is to design a low cost SSLEWS for risks reductions in the landslide prone areas of Rwanda. The system prototype was developed, tested, and deployed. The data are collected from the remote sites pushed to a web database hosted on a remote server where they are stored, analyzed, and visualized in a dashboard. Early warning messages are sent to the citizens around that area via SMS when the preset threshold values of the parameters inducing landslides are attained. In our future work, we shall integrate rainfall data and machine learning into SSEWS.

## References

1. Piciullo L, Calvello M, Cepeda JM (2018) Territorial early warning systems for rainfall-induced landslides. *Earth-Sci Rev* 179(April 2017):228–247. <https://doi.org/10.1016/j.earscirev.2018.02.013>
2. Piciullo L et al (2017) Definition and performance of a threshold-based regional early warning model for rainfall-induced landslides. *Landslides* 14(3):995–1008. <https://doi.org/10.1007/s10346-016-0750-2>
3. Early warning system | UNDRR. <https://www.undrr.org/terminology/early-warning-system>. Accessed 26 Jan 2021
4. Meteorological W, Wmo O (2017) Multi-hazard early warning systems : a checklist
5. Fathani TF, Karnawati D, Wilopo W (2016) An integrated methodology to develop a standard for landslide early warning systems, pp 2123–2135. <https://doi.org/10.5194/nhess-16-2123-2016>
6. Lacasse S (2008) Event tree analysis of Aknes rock slide hazard. In: 4th Canadian conference on geohazards: from causes to management, vol 2, no 2, pp 551–558
7. Silva BN, Khan M, Han K (2018) Internet of things: a comprehensive review of enabling technologies, architecture, and challenges. *IETE Tech Rev* 35(2):205–220. <https://doi.org/10.1080/02564602.2016.1276416>
8. Giordan D, Wrzesniak A, Allasia P (2019) The Importance of a dedicated monitoring solution and communication strategy for an effective management of complex active landslides in urbanized areas. <https://doi.org/10.3390/su11040946>
9. Segoni S, Piciullo L, Gariano SL (2018) Preface: landslide early warning systems: monitoring systems, rainfall thresholds, warning models, performance evaluation and risk perception. *Nat Hazards Earth Syst Sci* 18(12):3179–3186. <https://doi.org/10.5194/nhess-18-3179-2018>
10. Guzzetti F et al (2020) Geographical landslide early warning systems. *Earth-Sci Rev* 200(2019):102973. <https://doi.org/10.1016/j.earscirev.2019.102973>
11. Nadim F, Intrieri E (2011) Early warning systems for landslides: challenges and new monitoring technologies, pp 1–15
12. Wicki A, Stähli M (2021) Performance analysis of regional landslide early warning based on soil moisture simulations
13. A new perspective on the spatio-temporal variability of soil moisture\_stemporal dynamics versus time-invariant contributions\_2012.pdf
14. Wicki A, Lehmann P, Hauck C, Seneviratne SI, Waldner P, Stähli M (2020) Assessing the potential of soil moisture measurements for regional landslide early warning 2015:1881–1896. <https://doi.org/10.1007/s10346-020-01400-y>
15. Fensholt R, Abdelkerim AI, Eusuf MMRS (2016) Study of soil moisture sensor for landslide early warning system: experiment in laboratory scale sstudy of soil moisture sensor for landslide early warning system: experiment in laboratory scale. <https://doi.org/10.1088/1742-6596/739/1/012034>
16. Posner AJ, Georgakakos KP (2015) Soil moisture and precipitation thresholds for real-time landslide prediction in El Salvador. *Landslides* 12(6):1179–1196. <https://doi.org/10.1007/s10346-015-0618-x>
17. Kuradusenge M (2021) Rainfall and soil moisture data. [https://aceiot.ur.ac.rw/IMG/xlsx/sm-rf-data\\_allsites\\_martink.xlsx](https://aceiot.ur.ac.rw/IMG/xlsx/sm-rf-data_allsites_martink.xlsx)



# Decoding Visual Covert Attention Shift from EEG for Use in BCI



Swati Aggarwal, Nupur Chugh, and Arnav Balyan

**Abstract** The brain computer interface (BCI) is a communication and command pathway that allows the human brain to interact with an external device and produce output without relying on peripheral nerves and muscle movements. The primary goal of interpreting brain waves for brain computer interface (BCI) control in the visual field is covert attention to spatial and color characteristics. Here authors applied convolutional neural network to decipher the covert shift based on publically available data collected from 18 subjects during covert visual task. The results obtained declared that CNN outperformed among the conventional machine learning algorithms and obtained the promising accuracy in decoding covert attention shift. The efficiency of proposed method is furthermore compared with gradient boosting and support vector machine (SVM). The methods proposed in this paper have decoding performance that suggests they could be used in a brain–computer interface (BCI) framework based on visual attention shift.

**Keywords** Electroencephalogram (EEG) · Machine learning algorithms · Visual attention · Brain computer interface

## 1 Introduction

Brain computer interface (BCI) systems provide humans with a behavioral and covert means of transmitting a control signal to a computer. The primary objective of BCI research is to create a method that allows individuals to use brain signals to communicate with their environment without involving the motor cortex or utilizing any muscle activity. BCI has a broad spectrum of applications, including psychology, psychiatry, neurorehabilitation, and neuroscience.

Electroencephalography (EEG) has been the most common option for BCI applications among the currently available neuro imaging modalities due to its non-invasiveness, mobility, and low cost [1]. The electroencephalogram (EEG) measures

---

S. Aggarwal (✉) · N. Chugh · A. Balyan  
Netaji Subhas Institute of Technology, New Delhi, India

the activity of the brain that accumulates at a high rate in real time via the spontaneous electrical potential of the scalp [2]. This high temporal resolution of EEG oscillations captures some brain neural patterns that can be used to decode the intent of users.

Because of its robustness, the visual BCI dependent on external stimuli is gaining popularity and is a new research area in this field. The P300 event-related potential (ERP) and steady-state visual-evoked potential (SSVEP) are the most widely used triggers for the visual BCI [3, 4].

Most attention-related visual BCIs use oddball paradigm variations to discern the selected stimulus by focusing on time function. The process of selecting a stimulus is time-consuming and requires a significant amount of mind's endeavor. Spatial focus, which is an especially possible solution for impaired eye movement, is an alternate brain mechanism sufficient for BCI regulation. That is, without the need to shift the eyes, the covert spatial attention shift can be assigned to visual field objects [5]. Importantly, multiple target alternatives may be contained simultaneously in a single stimulus, allowing for greater performance with a correspondingly less trials.

Spatial attention to colored stimuli presented peripherally has recently been shown to enable precise, gaze-independent control of a four-class BCI [6]. The fact that paying attention to basic characteristics like color produces evoked potential that varies depending on the visual field in which it appears, recommend adapting N2pc for gaze-independent BCI. However, the use of covert spatial shift for BCI control has not been thoroughly checked [6]. The extraction of efficient characteristics to produce increased efficiency is a big and important aspect in EEG-based BCI systems, therefore developing acceptable approaches can result in the improvement of such systems. However, the potential benefits of exploiting variations in spatial focus in BCI operation have yet to be comprehensively explored [6].

Decoding EEG data provides useful information on brain mechanisms in terms of cognitive tasks [7–9]. Machine learning approaches have been discovered to be useful in processing brain inputs in order to extract brain properties that are discriminative. Deep learning, a subfield of machine learning, has recently acquired interest among researchers as a way for automatically extracting features to maximize class discrimination [10]. Deep learning systems such as convolutional neural networks (CNN) and recurrent neural networks (RNN) have seen a surge in popularity and success as a result of recent developments in machine learning. VGG-16 [11] and Residual-Net [12] are two CNN frameworks that have been introduced. A directed graph is generated among computational units along a sequence in a recurrent neural network, which is commonly utilized for time series data analysis. Although CNN and RNN have had a lot of success, they have not been thoroughly investigated for deciphering covert visual attention using EEG signals [13–16].

Deep learning could be a good way to decode covert spatial shift. It contributes to the creation of a solid framework that outperforms previous methods. To decode the accuracy of covert visual attention accuracy using time-varying EEG data, the authors use a CNN-based approach.

## 2 Methods

This section outlines about the dataset and the experimental setup used for acquiring the dataset used in this study. It further details about the methodology used for the decoding of spatial shift using convolutional neural network. The following are the steps taken for the proposed solution, as shown in Fig. 1.

### 2.1 Data Acquisition

The authors analyzed the data collection of 18 participants’ high-density EEG recordings. There was regular or corrective vision in all participants, and no neurological impairment was found. The data used in the experiments is taken from BNCI Horizon 2020 [17].

Participants were presented with a sequence of ten visual stimuli in which a red ‘x’-cross and a green ‘+’-cross were presented simultaneously in the opposite visual hemispheres. In the order of presentation, each shape was pseudo-randomized. Participants were asked to compare the green cross with the word ‘yes’ and the red cross with the word ‘no’ when reacting to questions and comments that were shown on the screen before the stimulus sequence presentation began. They only conveyed their response by concentrating their attention on the respective cross, while fixing their visual gaze at the center of the screen on a cross.

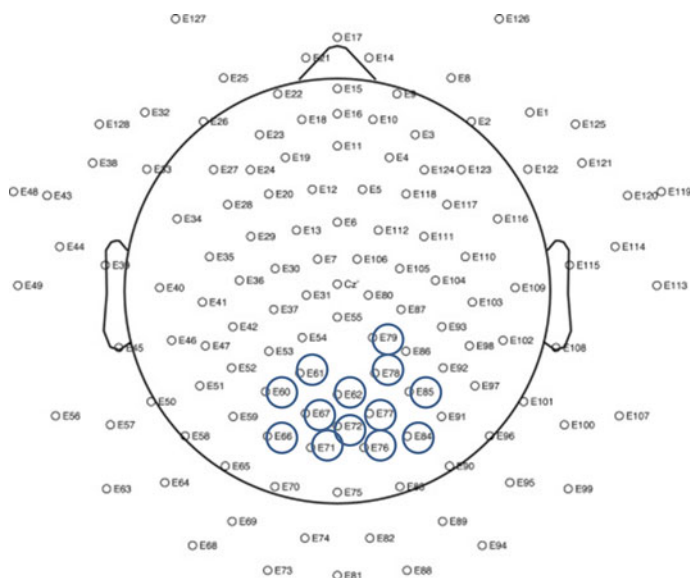
### 2.2 Preprocessing of Data

This section gives details of magnitude frequency responses and filters chosen for EEG data recordings.

A finite-impulse-response bandpass filter with 0.1 and 50 Hz cut-off frequencies for the first bandpass filtered the signals. Then the EEG recordings were down-sampled to 50 Hz. Channels corresponding to electrodes mounted over the occipital cortex circled in blue in Fig. 2 were used by the authors in this study. The data epoch was generated from the start of the stimulus to 750 ms after the start of the stimulus. The data epoch was generated from the start of the stimulus to 750 ms after the start of the stimulus. Epochs with  $y_i = 1$  are labeled with a green symbol on the left and



**Fig. 1** Proposed methodology for decoding covert spatial shift



**Fig. 2** Positions of EEG electrodes according to 10–20 system. The electrodes marked in blue are used in this study

a red symbol on the right. All the preprocessing of the EEG signals were done using EEGLAB [18].

### 2.3 Convolutional Neural Network (CNN)

Convolutional neural network (CNN) is made of neurons that have learnable weights and biases. Each neuron carries out dot product on input taken and optionally executes a nonlinearity after that. The weights of neurons regulate the influence among neurons. Neurons make up the layers of CNN that are organized in three dimensions: the input's spatial dimensionality (height and breadth) and depth. The depth refers to the third dimension of activation volume. CNN is composed of layers such as convolutional layers, pooling layers, and fully connected layers. CNN has less complexity and requires less memory since the weights of convolution layers are blended with fully connected layers during classification. CNN has been successfully employed for machine vision and medical image recognition, according to the literature [19–21]. As a result, the authors investigated if CNN might be used to decode the hidden spatial shift to colored symbol.

**Table 1** Summary of layers of proposed CNN architecture

Layer	Type	Shape	Parameter
conv2d	Convolution	(36,14,32)	320
dropout	Dropout	(36,14,32)	0
Conv2d_1	Convolution	(34,12,32)	9248
dropout_1	Dropout	(34,12,32)	0
maxpooling_2d	Maxpooling	(17,6,32)	0
dropout_2	Dropout	(17,6,32)	0
flatten		3264	0
dense	Dense	10	32,650
dropout_3	Dropout	10	0
dense	Dense	2	22

2.4 Proposed Architecture

The suggested model is made up of two convolutional layer blocks, which are followed by a maxpooling layer and a dense layer. A dropout strategy was used to apply penalty to the loss function by randomly eliminating inputs during the training to avoid the overfitting problem. After each convolutional layer, a dropout layer was inserted at a certain rate. The two CNN blocks’ outputs are combined in a dense layer with a softmax activation.

The summary of layers of proposed architecture is shown in Table 1. The binary cross-entropy loss function and the RMSProp optimizer are used in this study. The batch size and learning rate are set to 256 and 0.001, respectively, during training. To avoid overfitting, the authors execute 150 training iterations (epochs) and use validation stopping.

2.5 Conventional Method

The authors implemented canonical correlation analysis for extracting frequency domain features, and this method has successfully applied to EEG signals for SSVEP paradigm. The extracted features are then classified using conventional machine learning methods.

**Canonical Correlation Analysis (CCA).** Canonical correlation analysis (CCA) [22] is a statistical method for determining the correlation between multidimensional variables of a dataset and is widely used for frequency detection of SSVEPs [23, 24]. In this paper, authors used two set of signals of same dimension for analyzing the frequency component, one is EEG signals and the other is stimulus signal.

EEG signal is denoted by  $X$ , that was generated from the occipital cortex channels and stimulation signals  $Y_i (i = 1, 2, 3, \dots, N)$  denotes the reference signals which was created as

$$Y_n = \begin{bmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \\ \sin(4\pi ft) \\ \cos(4\pi ft) \\ \dots \\ \dots \\ \sin(2\pi nft) \\ \cos(2\pi nft) \end{bmatrix} \quad (1)$$

where  $f$  sampling frequency.

Since the given dataset was continuous in nature, to apply classification algorithms, 'Y' was converted to labeled data by splitting it into classes around the mean while  $X_{\text{train}}$  and  $X_{\text{test}}$  were normalized with the standard Min–Max Scaler. CCA thus used to transform the  $X_{\text{train}}$  and  $Y_{\text{train}}$  so created and this resulted in a maximized feature vector. On applying canocorr function, in each iteration of algorithm, there is decrease in canonical correlation value, thus authors used values that have level of significance  $p < 0.1$ . Data was transformed into 147 components. The correlation values obtained by CCA were used to equate the accuracy for decoding covert spatial shift. The authors have selected to perform experiments with support vector machine (SVM), widely used for EEG analysis and gradient boost. Both the classifiers used cross validation grid to hyper tune the parameters to achieve best score.

### 3 Experimental Results

#### 3.1 Performance Evaluation

The authors trained and evaluated the proposed network using EEG data from each person in this paragraph. There were 168 questions in each subject, with 10 trials for each question. As a result, each person had 1680 trials in total. Seventy percent were utilized for training (1176 trials), ten percent for validation (168 trials), and twenty percent (336 trials for testing) were utilized for testing.

The consequences of decoding visual attention have been demonstrated in terms of many performance metrics, such as precision, recall, and accuracy. The CNN network was used to analyze efficiency. Furthermore, it is compared with SVM, one of the widely used classifier for EEG analysis and gradient boosting.

**Table 2** CNN parameter setting

Parameter	Value
Convolution layers	2
Input size	36
Batch size	256
Optimizer	RMSprop
Epoch	150

**Table 3** Result of evaluation metrics on class-based samples

Classifier	Precision	Recall	f1score
Class 0	0.67	0.74	0.70
Class 1	0.73	0.65	0.69

During the experiment, the back-propagation algorithm trains each of the models, and the mini-batch gradient descent algorithm determines the gradient. In Table 2, relevant parameter settings for CNN are shown.

The precision, recall, and f1 score for class 0 and class 1 are shown in Table 3, illustrating the CNN network’s ability to detect spatial change in EEG time series signals. The class 0 and class 1 are defined as:

Class 0: (If answer was yes and positive target was on the left) or (if the answer was no and negative target was on the right).

Class 1: (If answer was yes and positive target was on the right) or (if the answer was no and negative target was on the left).

The author also plots the average accuracy and average loss for training and testing as shown in Fig. 3.

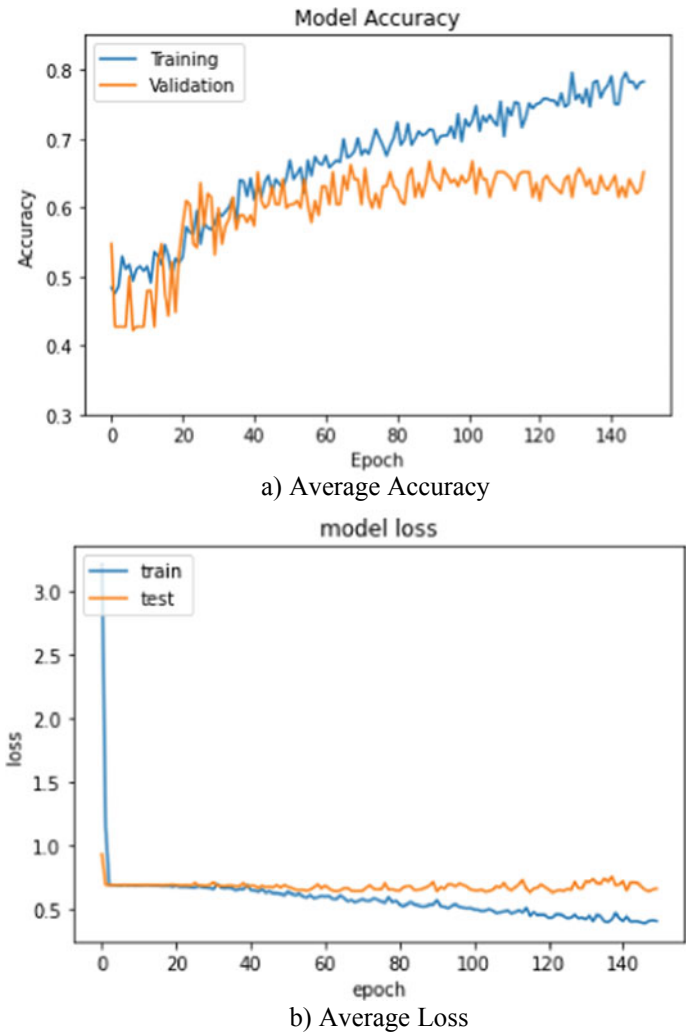
**3.2 Comparison with Conventional Methods**

This section describes the comparison result of CNN with conventional classifiers (i.e., SVM and gradient boosting). The performance measures accuracy and area under curve (AUC) on test datasets were determined and used for comparison. The accuracy percentage is the percentage of feedback that are correctly identified. At varying thresholds, the trade-off between sensitivity and specificity is represented by the receiver operating characteristic (ROC) curve. The area under the curve (AUC) determines the classifier’s aptness to discriminate between positive and negative classes [25, 26]. Table 4 shows how the proposed strategy outperforms established methodology pertaining to decoding accuracy and AUC across subjects.

From Table 4, as can be shown, CNN has outperformed among all classifiers resulted in highest decoding accuracy. Figure 4 depicting ROC curve of the classifiers further demonstrated that the result of the CNN is significantly finer than SVM and

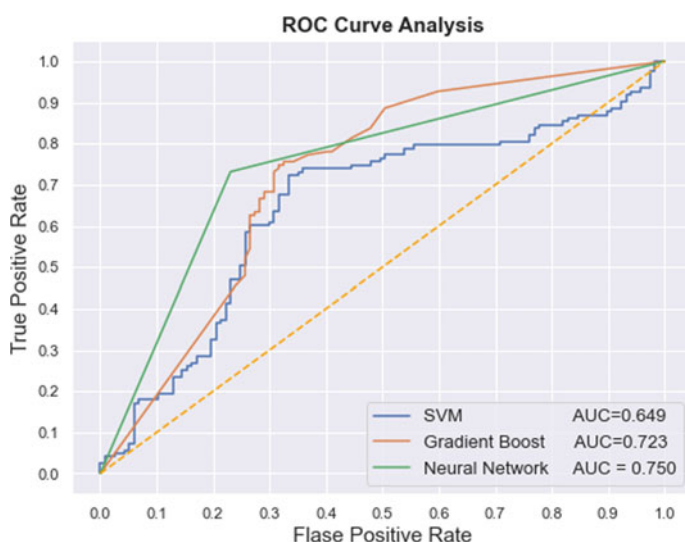
**Table 4** The accuracy and area under the curve (AUC) of the classifiers

Classifier	Accuracy (%)	AUC
SVM	66.6	0.649
Gradient boosting	67.9	0.723
Proposed method (CNN)	70	0.750



**Fig. 3** Plots for average accuracy and average loss for training and testing





**Fig. 4** ROC curve for different classifiers

gradient boosting for decoding covert visual attention than the previous frequency-based features.

As a result, the proposed approach is a good way for people with limited eye control to communicate by answering yes/no questions by concentrating spatial attention on a colored symbol.

## 4 Conclusion

In this paper, the practicality of the proposed methodology is demonstrated for decoding binary decisions through spatial attention focus on colored symbol using CNN. This procedure may lead to the development of BCI as assistive communicating tool for patients with serious motor neuron disorders. The authors have compared the proposed approach with conventional methods using frequency domain features. The results showed that the CNN has achieved the highest decoding accuracy toward covert attention shift.

This technology can give an instinctive BCI to elucidate EEG signals, as well as ameliorate interactions with persons who are unable to communicate through traditional ways for a variety of reasons. Furthermore, strategy may theoretically act as an assistive communication mechanism for individuals with serious eye movement impairments.

## References

1. Abiri R, Borhani S, Sellers EW, Jiang Y, Zhao X (2019) A comprehensive review of EEG-based brain–computer interface paradigms. *J Neural Eng* 16(1):011001
2. Malmivuo J, Plonsey R (1995) Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. Oxford University Press, USA
3. Fazel-Rezai R, Allison BZ, Guger C, Sellers EW, Kleih SC, Kübler A (2012) P300 brain computer interface: current challenges and emerging trends. *Front Neuroeng* (5):14
4. Liu Y-H, Wang S-H, Hu M-R (2016) A self-paced P300 healthcare brain-computer interface system with SSVEP-based switching control and kernel FDA+ SVM-based detector. *Appl Sci* 6(5):142
5. Posner MI, Charles RS, Brian JD (1980) Attention and the detection of signals. *J Exp Psychol General* 109(2)
6. Reichert C, Tellez Ceja IF, Sweeney-Reed CM, Heinze HJ, Hinrichs H, Dürschmid S (2020) Impact of stimulus features on the performance of a gaze-independent brain-computer interface based on covert spatial attention shifts. *Front Neurosci* 14
7. Jafakesh S, Jahromy FZ, Daliri MR (2016) Decoding of object categories from brain signals using cross frequency coupling methods. *Biomed Signal Process Control* 27:60–67
8. Gu Y, Farina D, Murguialday AR, Dremstrup K, Montoya P, Birbaumer N (2009) Offline identification of imagined speed of wrist movements in paralyzed ALS patients from single-trial EEG. *Front Neurosci* 3:3
9. Rafiee J, Rafiee MA, Prause N, Schoen MP (2011) Wavelet basis functions in biomedical signal processing. *Expert Syst Appl* 38(5):6190–6201
10. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
11. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
12. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
13. Andersson P, Ramsey NF, Raemaekers M, Viergever MA, Pluim JP (2012) Realtime decoding of the direction of covert visuospatial attention. *J Neural Eng* 9(4):045004
14. Ahmadi A, Davoudi S, Behroozi M, Daliri MR (2020) Decoding covert visual attention based on phase transfer entropy. *Physiol Behav* 112932
15. Davoudi S, Ahmadi A, Daliri MR (2020) Frequency–amplitude coupling: a new approach for decoding of attended features in covert visual attention task. *Neural Comput Appl* 1–16
16. Zhang D, Maye A, Gao X, Hong B, Engel AK, Gao S (2010) An independent brain–computer interface using covert non-spatial visual selective attention. *J Neural Eng* 7(1):016010
17. BNCI Horizon Homepage. <http://bnci-horizon-2020.eu/database/data-sets>. Last Accessed 31 Mar2021
18. Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134(1):9–21
19. Blanco G et al (2020) A superpixel-driven deep learning approach for the analysis of dermatological wounds. *Comput Methods Programs Biomed* 183
20. Ebrahimighahnavieh MA, Chiong R (2019) Deep learning to detect alzheimer’s disease from neuroimaging: a systematic literature review. *Comput Methods Programs Biomed*
21. Shahin A, Guo Y, Amin KM, Sharawi AA (2017) White blood cells identification system based on convolutional deep neural learning networks. *Comput Methods Programs Biomed*
22. Anderson TW (1958) An introduction to multivariate statistical analysis, vol 2. Wiley, New York
23. Lin Z, Zhang C, Wu W, Gao (2006) X Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Trans Biomed Eng* 53(12)
24. Bin G, Gao X, Yan Z, Hong B, Gao S (2009) An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. *J Neural Eng* 6(4):046002

25. Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* 27:861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
26. Hajian-Tilaki K (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp J Intern Med* 4:627–635. <https://doi.org/10.1017/CBO9781107415324.004>

# Soft Embedded Cores for FPGA



Praveen Kalkundri, Hansraj Guhilot, and Kalkundri Ravi

**Abstract** This paper presents different cores with user logic in field programmable gate array (FPGA). Different processor cores can be designed for different applications. These cores are used for complex electronics designs and prototyping different applications for FPGA as reusing of processors gives good time to market constraints. It provides different practical solutions for System-on-Chip (SoC) designs, hardware-software co-design, embedded applications, and FPGA designs. The paper consists of different softcore and their comparison in implementing on reconfigurable FPGA development environments. It helps to speed up the stages of description, synthesis, and designing for the user systems. The result discusses the comparisons of different embedded cores and their design methodology, design process, and specifications. It also emphasizes on different strengths these softcores provide for critical and safety applications.

**Keywords** FPGA · Softcores · MicroBlaze · S1 · LEON · OpenRISC · NIOS · Embedded system · Flexibility · Rapid prototyping · Programmable system-on-chip · IP Core

## 1 Introduction

Field programmable gate arrays (FPGAs) are a very popular choice for paradigm and improvement of products in a minute and reasonable quantity. An FPGA is a kind of Programmable Logic Device (PLD) that is used mostly for designing a general digital circuit that is limited to its circuit sizes. Any CAD (Computer-Aided Design) tool can be used for circuit design on FPGAs. Usually, an HDL (Hardware Description

---

P. Kalkundri (✉) · K. Ravi  
Gogte Institute of Technology, Udyambag, Belgaum, Karnataka 590008, India  
e-mail: [pukalkundri@git.edu](mailto:pukalkundri@git.edu)

K. Ravi  
e-mail: [rukalkundri@git.edu](mailto:rukalkundri@git.edu)

H. Guhilot  
K. C. College of Engineering and Management Studies, Mumbai, India

Language) softcore can be a microprocessor or microcontroller that can be fully customized, which can be synthesized on FPGAs, which are reprogrammable. A softcore processor targeting FPGAs is very flexible and easy to change. Hence it is easy and feasible to change its design parameters at any given time.

An Embedded system in an FPGA will mostly need some form of “controller” in its system. The controller can be a fully fledged microprocessor running an operating system or a simple microcontroller. Off-the-shelf is already available within the market and can be freely purchased and be used in a system. Sometimes an OTS processor-based approach will not meet our requirements [1], for example, an application that requires the functionality of the peripheral is not present in a discrete manner or where a board’s real estate is limited.

Parallel algorithm implementation is possible with devices like FPGA (field programmable gate array). Applications where sequential algorithms don’t require heavy processing or computation can be easily implemented on a microprocessor [2]. Many applications that need the advantages of both microprocessor/microcontroller with an FPGA array can use softcores implemented on FPGA. A softcore like MicroBlaze is a RISC Harvard architecture implemented on an FPGA chip [3], where power consumption is very closely monitored.

FPGA-based system-on-chip (SoC) can be designed by using Soft CPU cores. In such soft processors, the controlling work of the circuit is done by the CPU core which also supports random calculations and the rest of the design is accountable for parallel processing and interfacing.

Applications, where critical performances are required, are implemented in hardware. Since various systems tend to use common algorithms, reuse is much needed. Intellectual property (IP) cores are reusable components. An IP core can be used to build a system that is complex and dynamic in its design. IP cores are broadly divided into three categories like firmcores, hardcores, and softcores which are calibrated depending on their optimization, development, and adaptability of reuse. A softcore is synthesizable with HDL specifications that can be targeted to various devices.

The remaining of the paper is organized as follows: Sect. 2 survey of various softcore processors that are widely used. Section 3 gives a brief introduction to OpenSPARC S1 Core and its features. Section 4 gives a brief idea of The LEON3 processor and its configurations. Section 5 gives the details and description of the OpenRISC family and OR1200 IP core Architecture. Section 6 gives a brief idea about Nios II Processor Core. Section 7 describes Xilinx’s FPGA-based MicroBlaze soft processors. We draw conclusions and results in Sects 8 and 9.

## 2 Softcore Processors

The soft processor supports several other technologies that are present and enables it for a cost-effective softcore processor [4]. The key elements for the softcore processor were high-density FPGA. The HDL development tools environments are expected

to code complex processors and provide a low cost for the design. The first soft processor, Nios core, was released by the San Jose company and Altera in the year 2001.

The soft processor is mostly designed using an HDL, which is either Verilog or VHDL that is before synthesizing into a digital image before downloading the design into the device, which defines the logic build within the FPGA as per the required design. The soft design consists of all elements required for the design of the processor. The processor consisting of memory controllers, ALU (Arithmetic Logic Unit), interrupt controllers, bus controllers, and caches are all included in the synthesized bitstream [4]. The major advantages the soft CPU provides are its flexibility and low cost and reusable and also the added ease of building multiprocessor systems. The soft processor-based which requires hardware solution within an FPGA design can be updated, improved, modify or debugged without any physical hardware changes [4]. The only limitation of and FPGA devices is the amount of logic fabric and memory, and the maximum speed it is capable of operating.

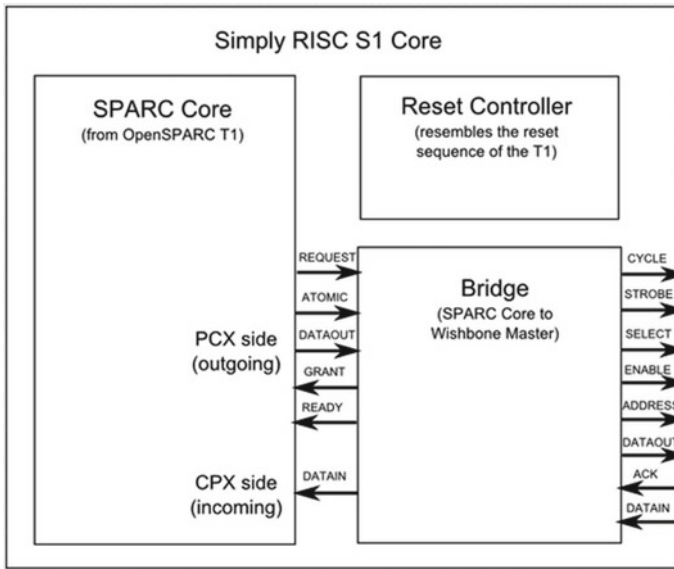
Presently many soft processors are available for different devices like FPGA. The old microprocessors which are present but are difficult to replace because of their wide use such as 8088, Z80, 6502, and 6809 are typical processors available as free softcores [4]. Many new custom processors are presently available such as TSK3000, Pico Blaze [5], ARM Cortex, MicroBlaze, and the legacy Nios and Nios II are also used for FPGA's and third-party providers. There are other high-performance softcores such as OpenRISC, LEON3, and OpenSPARC T1 by Sun Microsystems. The number of softcores and their performance and their speed through optimization are being improved.

### 3 S1 Core

The S1 Core is a simple RISC and a 64-bit CPU Core. S1 core is based upon the OpenSPARC T1 microprocessor. It is the first design which is released by Simply RISC and a cut-down version of the OpenSPARC T1 processor (see Fig. 1). The major difference between T1 and S1 core is S1 consisted of a 64-bit SPARC Core whereas T1 had 4, 6, or 8 cores. S1 core consists of a basic switch controller and a reset controller too. It also has a Wishbone bridge which is a standard and extremely easy bus for data exchange between IP cores. S1 core can be executed on an  $\times 86$  Linux machine by using free tools.

S1 Core was designed by Simply RISC which is licensed under GNU General Public License by Sun. S1 core is an open-source hardware microprocessor and also a decreased version of UltraSPARC T1, Sun Microsystems [6].

SPARC is a RISC instruction set architecture (ISA) that was introduced by Sun in mid-1987. SPARC design was based on scalable processor architecture. SPARC is a 32-bit-based architecture that was preferred and used in Sun's Sun-4 server systems and workstations. Sun-4 was a substitute for Sun-3 which had Motorola 68,000 family processor. Sun-4 was mostly used in server systems and workstations.



**Fig. 1** An internal SPARC S1 core diagram

Further, the SPARC was extended from 32-bit to a 64-bit processor. These 64-bit Sun-4 processors were used in servers manufactured by Fujitsu, Solbourne, and many other companies. SPARC is a CPU having an instruction set architecture (ISA) which is obtained from a reduced instruction set computer (RISC) lineage.

SPARC V9 includes the following principal features [6, 7]:

- Has 160 general purpose registers and can have 72–640 general 64-bit registers.
- 64-bit virtual addresses, 64-bit integer data and 9 added registers and 16 more dual precision.
- It has improved system performance with additional registers added which reduces the memory traffic. The SPARC V9 has 16 new registers with double-precision floating-point.

SPARC V9 includes the following principal features [6, 7]:

- Has 160 general purpose registers and can have 72–640 general 64-bit registers.
- 64-bit virtual addresses, 64-bit integer data and 9 added registers and 16 more dual precision.
- It has improved system performance with additional registers added which reduces the memory traffic. The SPARC V9 has 16 new registers with double-precision floating-point.
- It can launch many instructions in a single given time, hence achieving more parallelism. It has big-endian instructions.

Sun released a new version of UltraSPARC Architecture 2007 in the year 2007 called the UltraSPARC T2. In August 2012 an updated version with new specifications called the Oracle SPARC Architecture 2011 was released by Oracle Corporation. Further in 2015 Oracle released a new version again called SPARC M7. Sun's SuperSPARC and UltraSPARC-I were the most accepted processor among so many various implementations of SPARC.

## 4 LEON

LEON is a 32-bit RISC architecture and an instruction set microprocessor core which is rooted from SPARC-V8[8] synthesizable using VHDL. The European Space Research and Technology Centre (ESTEC) released the LEON architecture which was a part of the European Space Agency (ESA), after a short-term association with Gaisler Research.

LEON processors are programmable logic that can be implemented into an ASIC or an FPGA. The LEON processor series was similar to the SPARC-V8 RISC architecture. The older versions of LEON2 (FT) had a 5-stage pipeline whereas the latter versions were updated to a 7-stage pipeline. By using the specific graphical configuration tool, the LEON2 and LEON2 (FT) system-on-chips (SoCs) can be reconfigured and reused based upon the design requirements [1].

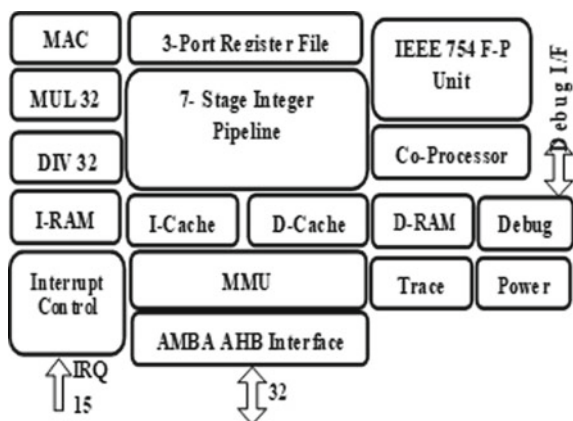
The 32-bit LEON3 is derived from SPARC V8 architecture which is a synthesizable VHDL model developed by Aero flex Gaisler AB in Sweden [9]. The processor model can be an arrangement by its functional units according to their nature, number, chief characteristic, and most preferable for SOC designs [10]. The source codes can be used for research and education purpose as it is permitted under GNU GPL license.

The few features of the LEON3 processor that can be highlighted are listed below [11]:

- The processor has an advantage which consists of hardware divide, multipliers, and MAC units.
- It has Harvard architecture with snooping, i.e., separate instruction and data cache.
- Extensively configurable.
- It has fully pipelined IEEE-754 FPU which provides high performance with 1.4 DMIPS/MHz and 1.8 Core-Mark/ MHz
- The caches can be configurable into 4 different ways and 1–256 Kbytes/way with Least Recently Used (LRU) or Least Recently Replaced (LRR).
- It has 1–512 Kbytes of data scratch pad RAM and local instruction.
- It has a clock gating power-down mode and a single-edge clock.
- The LEON3 processor includes an advanced pipeline as long as 7 stages.
- SPARC Reference MMU (SRMMU) with configurable TLB and AMBA-2.0 AHB bus interface
- It is available with on-chip debugging support.
- Symmetric multiprocessor support (SMP).



**Fig. 2** Structure of the LEON3 processor



- AMBA -2.0 Advanced High-performance Bus (AHB) interface.

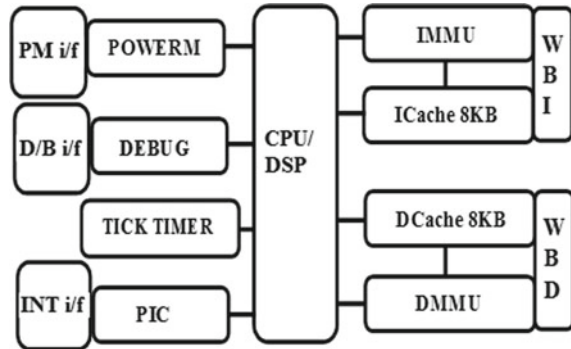
LEON3 is standardized to the IEEE-1754 (SPARC V8 instruction set with V8e extensions) architecture and mostly used for embedded applications having both low complexity and high performance with less power utilization. The LEON3 core (see Fig. 2) has a Harvard architecture with the 7-stage pipeline, separate data caches, instruction, hardware divider/multiplier, on-chip multiprocessor extensions, and debug support. The LEON3 processor is highly programmable. It is suitable for high real application as LEON3 processor is available with the fault-tolerant version which gives total immunity from radiation effects [12].

## 5 OpenRISC 1000/1200

OpenRISC 1000 is a 32/64-bit RISC architecture that belongs to a family that is freely available and has an open-source RISC-based processor core [13]. It supports many applications as it can be implemented on a wide range of chips and systems. The architecture can be configured for designs with less complexity, power optimization requirement, emphasis on performance, scalability, and flexibility. OpenRISC 1000 architecture is mostly used in applications where medium and high performance embedded is required.

The OR1200 is a RISC-based Harvard micro-architecture with 32-bit virtual memory support (MMU) and 5 stage integer pipelines. The default cache has a 16-byte line size which has a 1-way direct-mapped 8 KB instruction data cache that is physically tagged. It also comprises a debug unit, high-resolution tick timer, real-time debugging, power management support, and programmable interrupt controller [14]. OR1200 in default mode consists of 1 M transistor. OR1200 mainly focuses on embedded and networking applications. The OR1200 IP core can be configured by the user [15] and has a high-performance cache and MMU subsystems.

**Fig. 3** OR1200 IP core architecture



A generic architecture of the OR1200 IP core is shown in the figure (see Fig. 3), and it consists of several building blocks:

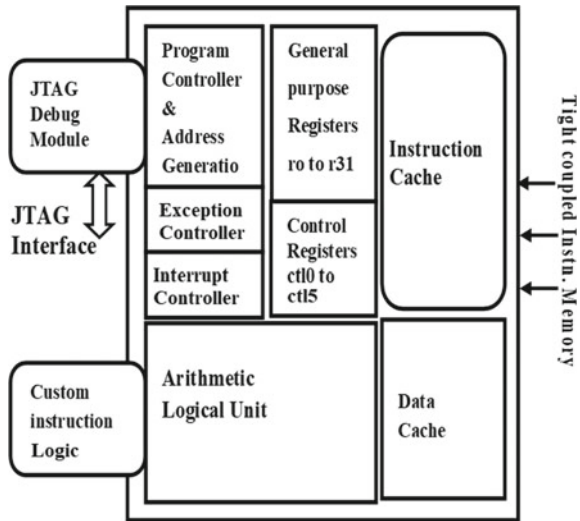
- Direct-mapped data cache and CPU/DSP central block.
- Data Memory Management Unit (MMU) is built on hash-based data translation look-aside Buffer (DTLB).
- The instruction cache is directly mapped and an interrupt controller.
- Power management unit for managing power functions and interface.
- It uses a tick timer which is a clocked RISC utilized by the operating system.
- It has debugged unit that assists software development to the system debug.
- It has two WISHBONE interfaces connecting the core to the external peripherals and memory subsystem.

## 6 Nios II

The Nios II is softcore embedded processor architecture which is a 32-bit RISC processor designed by Altera family [16]:

- It has a 32-bit instruction set and a general purpose register.
- Nios II has 32-bit external interrupt sources.
- It used Eclipse IDE and GNU C/C++ toolchain environments for software development.
- A  $32 \times 32$  multiply and divide is possible with a single instruction that gives the 32-bit result.
- It has 250 DMIPS and 32-bit address space and data path.
- It has a good hardware debugging module.
- All Nios II processor systems have consistent Instruction Set Architecture (ISA)
- A single precision-point operation has separate floating-point instruction.
- It is capable of using different on-chip peripherals and off-chip memories/peripherals.
- Its analysis data, instruction, and also signals on FPGA design in real time.

**Fig. 4** Nios II processor core block diagram



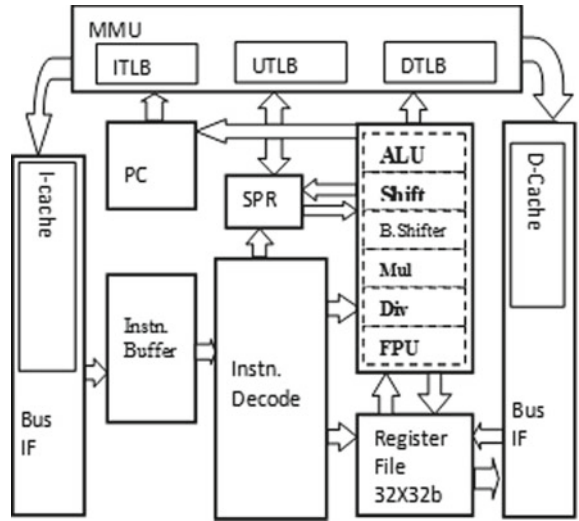
It has exclusive instructions for executing 128-bit and 64-bit multiplication. Figure (see Fig. 4) describes the hardware structure of the Nios II processor and the functional units of the Nios II architecture [17]. This architecture illustrates an instruction set architecture (ISA) which is one of the major programming abstractions [18]. The ISA has different functional units that are mending instructions. The processor core does not include the connection logic or its peripherals. Whereas it includes only the circuits required to build the Nios II architecture. The Nios II processor a customizable general purpose RISC processor core, unlike another fixed microcontroller. The features can be configured based on selective processor feature, peripheral, performance, and functions needed for applications that use any Intel FPGA family. Intel provides a free FPGA IP evaluation mode for Nios II processors without a license where the behavior of processors without systems can be observed.

## 7 MicroBlaze

MicroBlaze is a 32/64-bit soft processor based on RISC Harvard architecture designed by Xilinx. It has high configurable architecture options like Memory Management Unit (MMU), cache, AXI or PLB interface, Floating-Point unit (FPU), configurable pipeline depth [19], and many such features as shown in figure (see Fig. 5).

MicroBlaze provides a fast, reliable, and high user-configurable microcontroller or embedded processors in FPGAs [20]. It can also be used as a coprocessor for a few ARM SoCs. MicroBlaze provides 70 different configuration preferences as it can be extremely personalized [21]. To improve the size, MicroBlaze can be designed to

**Fig. 5** Functional block diagram of the MicroBlaze core



**Table 1** Comparison of IP cores

CPU Core	Architecture	Clock Freq in MHz	Bits	Pipeline Depth	License
OpenRISC 1200	OpenRISC 1000	50	32	5	Open-source (LGPL)
Leon	SPARC-V9	66	32	7,5	Open-source (GPL/LGPL)
MicroBlaze	MicroBlaze	200	32	3,5	Proprietary
S1 Core	SPARC-V9	250	64	6	Open-source (GPL)
Nios	Nios II	290	32	6	Proprietary

utilize 3-stage pipeline mode, and to optimize speed it used 5-stage pipeline mode [22]. Hence it delivers faster Dhrystone Million Instructions per Second (DMIPS) performance than any other softcore processor.

Xilinx provides an end-to-end solution consisting MicroBlaze embedded processor, which provides customized building embedded applications faster as it reduces the development time. Xilinx provides a personalized environment that provides a trade-off to achieve the best results required for embedded designs [23].

## 8 Results

From Table 1, it can be observed that Nios II and Microblaze are mainly implemented on FPGAs. Each of the cores has its specific performance characteristics and also are optimized for few critical applications. With the newer versions from Intel's Nios and Xilinx's make Microblaze, the throughput performance will be enhanced. These two processors are preferred as the royalty is free. Although both of these processors are synthesized and their development is locked to vendor specific FPGAs, a lot of IP blocks are also available which reduce the development time [24] and enhance the performance.

## 9 Conclusion

The MicroBlaze core can be configured as a microcontroller or as a real-time processor which has an on-chip memory running a FreeRTOS or also as an application processor which consists of a Linux-based memory management unit. On the other hand, the Intel Nios II softcore can be configured as either Nios II economy core or Nios II fast core. The Nios II economy core is suitable for microcontroller applications and the Nios II fast core provides wide and simple configuration options that support both real-time and application processing applications. The Nios II core is also a certified design for DO-254 compliance for military, aerospace, and avionics applications.

One of the major differences between MicroBlaze and Nios II is that Nios II provides up to 256 used defined instructions that increase the CPU instruction set and also boost the timing critical in software. It also helps in reducing the software inner loop which plays an important role in DSP applications or computational critical applications. Unlike Nios II, MicroBlaze is highly customizable and enables more than 70 user-configured options. It has few special instructions like Sleep, Hibernates, and Suspend Mode/Instructions suitable for power optimization. Most of the features and support for Nios II require licenses separately. This helps in selecting an appropriate processor core for its applications which optimizes the performance and validates the software development to provide a perfect combination of feature set, performance, area, and cost.

## References

1. LEON2 Processor User's Manual XST Edition. Gaisler Research (2005)
2. Guedria L, Hubaux D, Legat J-D. Soft-core FPGA processor based platform for embedded systems rapid prototyping, [White paper], Centre of excellence in information and communication technologies (CETIC), Available: [https://www.cetic.be/IMG/pdf/SoftCore\\_FPGA\\_Based\\_Platform\\_for\\_Rapid\\_Prototyping-2.pdf](https://www.cetic.be/IMG/pdf/SoftCore_FPGA_Based_Platform_for_Rapid_Prototyping-2.pdf)

3. Anderson SA (2013) Real-time embedded-four soft-core processors for embedded systems. Available: <https://www.design-reuse.com/articles/31120/soft-core-processors-for-embedded-systems.html>
4. AppleLogic, Processors, Article (1988). Available: <http://www.applelogic.org/Processors.html>
5. Praveen K, Guhilot H (2013) FPGA implementation of contactless position sensor using FSR (Force Sensing Resistor). IEEE International Conference, pp 1027–1031
6. SPARC international INC (1991) The SPARC Architecture manual USA. Version 8:4–65
7. Marcus E, Michael U, Rich D (2001) OpenRISC: Overview, Sept 2001
8. Aeroflex G, Website, LEON3 Processor, processor
9. Gaisler Research Website. (Online). Available: <https://www.gaisler.com/index.php/products/processors/leon3>
10. GRLIB IP Core User's Manual, Gaisler Research, Feb 2006
11. Actel, IP Docs, "LEON3", version 1.1, pp 1–10, March 2010
12. Gaisler Research (2014, Jan 18). LEON Series (Online). Available: [http://en.wikipedia.org/wiki/Leon\\_%28cpu%29](http://en.wikipedia.org/wiki/Leon_%28cpu%29)
13. Opencores.org, Website (2006). [www.opencores.org](http://www.opencores.org)
14. Lampret D (2001) OpenRISC1200 IP Core Specification
15. OpenRISC 1200 IP Core Specification (March 2001), (Online). Available: <http://www.openrisc.net/or1200-spec.html>
16. Intel Corporation, NIOS II processor reference handbook, October 2020
17. Intel, Website. (Online). Available: <https://www.intel.in/content/www/in/en/products/programmable/processor/nios-ii.html>
18. Altera Corporation (May 2011) (Online). Available: <http://www.scribd.com/doc/86472321/Nios-2-Reference>
19. MicroBlaze Data Manual, Xilinx embedded systems. (Online). Available: <http://www.xilinx.com/tools/feature/csi/microblaze.htm>
20. Seely J, Erusalagandi S, Bethurem J (2017) The microblaze soft processor: flexibility and performance for cost-sensitive embedded designs
21. Xilinx, Quick Start Guide: MicroBlaze Soft Processor for Vitis, 2019
22. Xilinx, Xilinx Microblaze processor: scalable performance and industry-leading flexibility, Copyright 2015–2018 Xilinx, Inc
23. Mhadhbi I, Rejeb N, Othmen SB, Saoud SB (2014) Performance evaluation of FPGA soft cores configurations case of Xilinx MicroBlaze. Int J Comput Sci Commun Inf Technol (CSCIT) 1:14–19
24. Yiannacouras P, Steffan JG, Rose J (2007) Exploration and customization of FPGA-based soft processors. IEEE Trans Comput Aided Des Integr Circuits Syst 26(2):266–277. <https://doi.org/10.1109/tcad.2006.887921>

# Social Media Enabled Rehabilitation Services: Influence of Covid-19 Metaphors



Sreejith Alathur 

**Abstract** This study analyses the social media discussions on disability during the Covid-19 in the Southern part of India. The purpose is to assess the e-participants' attitude toward disability problems during their social media participation. Participant observation and focus group discussions with citizens often post in social media disability groups. Existing studies less reported the social media disability discussions in regional language, even though it is vital to explore justice at the grassroots level of disability inclusion. The current findings show that citizens' ill-conceived understanding of disability challenges during the covid-19 results in an unsupportive social media environment. The extreme and exploitation expressions in regional language groups show the lack of local support from the common public for the disabled during the pandemic. The lack of non-institutional voluntary metaphors emphasizes the need to improve the disability support structures at the regional level. Recommendations to enhance e-participation competency are provided.

**Keywords** E-participation • Disability • Covid-19 • Social media • Metaphors • Competency

## 1 Introduction

Different kinds of digital platforms enable varying nature of e-participation requirements in a democracy system. The participants in these forums expect to possess the competency to generate attractive and meaningful content for policymaking. The knowledge, skill, and interpersonal influencing capacity [1] of the e-participants may vary in these forums. Existing network studies focus on key influencers and shaping social influence and point to factors like 'assortative, joint distribution of influence, and susceptibility' and fear of the manipulative capacity of platforms [2]. The mobilization or reinforcement nature of such platforms' contents by emphasizing the strength of ties is also studied [3]. In weak ties, the scarcity of local

---

S. Alathur (✉)

National Institute of Technology Karnataka, Mangalore, India

e-mail: [a.sreejith@iitdalumni.com](mailto:a.sreejith@iitdalumni.com)

knowledge resources demands the blends of local and global praxis for transnational influence, especially for the marginalized groups [4]. Disability theories are sophisticated by its' potential to enable universal design. Disability group is excluded mainly due to the misclassification of their identity in digital platforms. The e-participation studies yet to coalesce the medical and social model of disability. Paper discusses disability awareness and competition strategies. Thus, digital platforms' content and relationship are crucial for 'consciousness' that results in meaningful participation. This paper addresses the overlapping of e-participations in the public sphere and awareness among the e-participants.

The misclassification of disability identity in the digital platform excluded marginalized groups. The lack of disability awareness results in identity dialectics among the e-participants about the heterogeneous disabled groups. E-participants influence the awareness building; if it results in false consciousness, the platforms radicalize and get manipulated. For a meaningful e-participation to occur for the marginalized groups—persons with disabilities—the knowledge about the specific requirements pertinent to diverse groups of disabilities is essential. Such diverse groups' complexity is because popular impairment theories limit with expressions of medical requirements and are yet to address the current development practices of rehabilitation approaches from social aspects [5]. The e-participants must address the misidentity and disability problems competitively to shape the digital platforms for an enabling role.

The 'hierarchy' among disabled people, disabled studies, and between the 'temporarily abled bodies' [6] is considered a challenge in explaining the instances of disability and framing unified disability programs. The lack of an adequate level of 'novel metaphor' about disabilities and the presence of 'creative metaphor' with ableist expressions [7] force the e-participants to 'deconstruct' the knowledge about disabilities. Scholars [8] point to the transformation from the medical model of disability to the social terms that help to understand disability for a better civil society. Structured medical knowledge about impairment fails and social awareness struggle is due to political, religious, trans communities with disabilities, diversity in norms, and the nature of disabilities. Globally, the accessibility, lack of compliance in civil and political rights, etc., differ concerning persons with disabilities. The transnational efforts like the United Nations Convention on the Rights of Persons with Disabilities (UNCRPD) can be considered as efforts to blend the local and global understanding of the social model of disabilities. Sustainable development initiatives using Information and Communication Technologies (ICTs) help diffuse the knowledge and to attain equality for these marginalized groups. However, studies are not available in this direction. The current paper analyses the disability 'awareness, skill, and experience' of e-participants and depicts the competency level for making inclusive digital platforms.

Due to diversity and spatial distance, the researches on disability awareness emphasis for regional studies even though its' praxis is transnational [4, 9]. Analyzing citizens' behavior in the global pandemic, scholars identify the 'responses were local' due to regional influence [10]. Health emergencies often invoked medical and social constructs. The health emergency pandemic model's governance is calculated



through the global exclusion indices, including Stringency Index, Democracy index, Power-distance index, and Human Rights Index [11]. The World Health Organization's and local government agencies released disability-inclusive guidelines during a health emergency. However, civil society metaphorically mapped the lockdown and social distancing measures by comparing persons with disabilities' lifestyles. This shift to problematic categorization by conspiring networks and overriding the official knowledge sources is often reflected in social media during the pandemic. It may hammer the inclusive structure of such platforms. However, factors structuring social media categories is underreported. The research question is: What influences the e-participation competency? With this aim, the study carryout social media metaphoric analysis to understand disability awareness during the health emergency pandemic and the identity dialectics among the e-participant.

## 2 Theoretical Background

If the e-participant can empower the target groups to participate in a democracy effectively, it can be considered meaningful e-participation. E-participation studies reported the importance of improving the e-participation environment and demand competency building strategies in such platforms.

**Theories of meaningful e-participation:** The competency of e-participants is multifaceted, a prominent view of e-participation literature argues for the heterogeneous viewpoints, and they measure e-participation using ambiguity measures. While using ambiguity as a measure of e-participation effectiveness, the process reduces the reinforcement and mobilizes the citizens for more meaningful participation. Studies on conspiring networks point to the homogenous digital environment. Here, the e-participation leads to the development of a radical environment. The initiators struggle to enable meaningful participation due to the reinforcement in radical views among the citizens. The cause of radicalization can occur due to hate content, false information, which are part of the conspiring network. The research gap in this domain is yet to explain why the sensitization awareness initiatives limit is reducing the false consciousness.

Sensitization is commonly relaying the theories of power, and it expects to improve the perceptions, reorganization, and accommodating the 'others' by enhancing the understanding. While formal or informal mechanisms develop the awareness, a person engages in an action that influences his own life or the society he belongs [12].

In the initial level of e-participation, the e-enabling digital environment is provided by attractive content that stimulates silent digital citizens' belief. The non-participants' competency is limited to reception of available contents, and the perpetrators diffuse content for shaping an extreme and exploitation network. In the second level at the knowledge realm, reflections of self often develop within the e-participants. The implicit representation of the causality of oneself provides active participants in democratic forums. However, this group is less vulnerable than

the silent since the participation takes place with state of affairs knowledge with a fallacious conclusion about the truth in their knowledge they have experienced, irrespective of the totality or pastness of the situation. In the next competency level, the awareness provides an explicit representation of the structure and order of an e-participation system. Here, the thinkers internalize the group values and give the relationship of class-inclusion structure for the system. However, the group awareness in this system can also evolve into more than ‘basic human needs’ that are not necessarily peaceful, empathetic, secular, and conflict-free [13, 14].

### 3 Research Design

Scholars consider mapping metaphors to understand better the problems [15]. Scholars recommended the metaphorical analysis of cyber cultural studies, especially disability metaphors, to include marginalized groups [16]. However, metaphors that influence disability awareness among the e-participants are yet to explore in popular languages. The current study focuses on metaphors of cyberspace during the health emergency pandemic in Kerala, India. To assess the metaphors that may influence the disability awareness among the e-participants, the e-participation in Malayalam (native language of Kerala, India) during the health emergency pandemic is considered for data collection and analysis. Justifications for choosing this case are—(i) access to education and English proficiency of group members, including persons with disabilities in the study region (ii) disability awareness studies demand regional analysis, and (iii) The medical and social aspects of health emergency pandemic responses. Participant observation in various disability cyberspace activities helps identify social media groups—these groups show increasing membership and sharing meaningful disability-related content regularly. Social media groups pertinent to support persons with disabilities are considered for the study. Participant observation of native language posts related to a health emergency and persons with disabilities was also carried out. Focus group discussions with social workers, advocates, technologists through online forums until May 2021 is also considered. Thematic analysis was carried out using the data, and the disability metaphors are reported.

### 4 Results

During the participant observations of a health emergency, the author noted that the remarks on ‘disabilities’ are more during this pandemic, unlike other disasters. Health emergency preventive measures like lockdown are often mapped with the lifestyle of persons with disabilities. Remarks about the possible long-term disability due to ventilation among health emergency survivors [17] are also observed. The metaphors identified from the health emergency data is explained under various themes below.

**Solidarity:** metaphors reflecting the identity within disability groups were observed during the pandemic. Social network groups of persons with disabilities share solidarity about the types of disability the group members suffer or discuss. The political representation of the group members who run their disability services posted it as promotional content and often discouraged by group members, stop symbols—X—also put for the post in regional languages. A respondent told:

...my groups frequently receives comments like ‘stop such non-disability-related messages’..... ‘this group is for what’..... ‘some people are hearing impaired in this groups, hence do not put such messages as they cannot understand’... ‘heated arguments over language of post was also happened’ ....

The response shows members have various impairments and restraint from any cultural exchange. Even the groups with all types of disabilities are insisting on accessible content and the same language. A muscular dystrophy member said:

I only joined in this group....all my Facebook friends I got from this group...my only use of the Internet is to browse this group.

The members commonly report’ delay in getting their benefits, and this is often reported with comments that some members are not aware of such schemes themselves. A political party representative responded that:

such responses are common, we have conducted a number of welfare activities, and we don’t have any politics.... I should use this platform to discuss our service delivery... we don’t want to change the word ‘handicapped’ from our association name, as the establishment date is important for most benefits

It shows, more than disability activism, the group admins and representatives are motivated to build solidarity and group awareness to attain better living conditions.

**Quarantine and memes:** Preventive measures such as ‘lockdown’ and social distancing is metaphorically discussed as the lifestyle of persons with disabilities in social media. The e-participants frequently posted the receipts of penalty laid by the police officials for breaching the lockdown. Even as preventive measures, the lifestyle of ‘disabled’ is not acceptable for the abled bodies. However, other than the ‘medical condition,’ the e-participants who are disabled often posted about their proud achievements during the lockdown period. Thus, awareness developing around health emergency preventive measures with quarantine memes does not completely reflect the novel metaphor.

**Misdemeanor and Joke praxis:** Offensive videos were viral during the health emergency. In those videos,’ citizens breach the lockdown and to escape from the officials, they fake their body movements by imitating persons with disabilities. The social workers responded to the incidents with hashtags to ban social media and metaphors of ‘twisted humor’ and ‘a joke at our expense.’ The sadistic joke metaphor in videos reflects the false awareness about persons with disabilities. The videos in which animals are showing respect for traffic signals were also circulated during this period. This shows a severe identity dialectic about these marginalized groups.

The metaphors of representation to exclude the groups under the medical model are often reported during a health emergency.

Sickness: The exemption of persons with disabilities from emergency services during a pandemic is ‘not appropriately’ treated under ‘sick leave’ with historically banned expressions of ‘handicapped’ [18]. One disability right social worker responded:

Terrible ....Some granted Privilege Leave and & Special Leave but disabled discriminated...  
Ridiculous.

In the above incident during a health emergency, the metaphorical classification of a disabled preventive measure as ‘sickness’ was considered under the pregnancy scenario; it reflects the stigma of illness and integration of disability with the feminine. The possible loss of salary during a pandemic and the limited opportunities to avail preventive measures unless reformatory support was prevalent during a health emergency.

**Compulsive expressions:** Various documents that may be discriminatory toward persons with disabilities were shared during the health emergency. The compulsory retirements based on disability influence on performance were discussed during the pandemic in various disability forums. The participants point to the ableist expressions and the limited training that may result in the authorities disregarding law. The disability hate metaphors are shared by the ‘interferes’ and ‘burden’ terminologies to exclude the persons with disabilities and to discord the international guiding principles.

## 5 Discussions

The academic community often assesses knowledge about disability and its social constructs. Unfortunately, in developing nations, persons with severe disabilities less likely to attain education and not in the leisure of knowledge creations. The lack of experience makes the domain of knowledge mostly by observation-based. Due to limited resources and political agendas, the research funding often encourages considerable scope and high visibility studies. This lack of opportunities to avail long-term funding restrains the scholars from venturing into this disability domain and attains skills. Working with ‘traumatized individuals’ also leads to ‘secondary trauma.’ [19] The frustrated responses in disability groups make the limited regional reporting by scholars that to types of disabilities they are exposed to.

The desire to unify the concept of disability is evident from networks. Results show a cross representation of group members and the emergence of common disability discussion forums in developing nations. Even with diverse disabilities and cultural backgrounds, most groups are constraints to discussions of disability and life sufferings events of disabled members.

However, the quarantine memes with expectations of the life routine of disabled are not necessarily reflected complete understanding about the sufferings which are bed-ridden and restricted in movements. The ableist discussions on the lockdown period on cyberspace describe their cooking and family time. The poverty and routine

hygiene difficulties of persons with disabilities are mostly ignored, and the metaphors are dialectic to address the identity of persons with disabilities.

Plead for accessible contents in terms of regional language and sign language also points to the existing ownership identity based on diverse disability groups' e-participation competency. This supports [20] focus of the Internet that the metaphor 'cyberspace' is a property-based regime. The e-participation competency will be for those who trespass the accessible fencing.

**Implications for theory and practice:** Police authorities and various agencies adopt online trolls with native language movie scenes for awareness purposes during a health emergency. Such digital content was shared extensively by the citizens. However, the metaphors use of trolls for awareness initiatives related to persons with disabilities is less adequate in cyberspace. The health and rehabilitation agencies can adopt disability metaphors, including trolls, for inclusive cyber culture. To ensure cyberspace supports persons with disabilities during pandemic scenarios, regulatory measures and cyber cultural improvements need to occur.

Cyberspace theories less discussed the marginalized group and possibilities of universal design and transnational 'cultural exchange solutions' [21] by including theories of disabilities. The current study provides metaphorical disability representations that take place in developing nations. The metaphorical reviews help to explore existing identity dialects in the domain of disability studies. The rehabilitation strategies using cyberspace need to revamp by considering the problematic nature of cyberspace periodically.

## 6 Conclusion

E-participation research with disability awareness can be considered the third wave of disability writing were the last wave was during the 1990s. In India, right-based movements were propelled only after 2017s. The right-based movements are expected to improve awareness about persons with disabilities; however, the metaphors of social media contents during health emergency was dialectic and disturbing. The e-participants expressions about persons with disabilities were limited by a lack of knowledge, skill, and experience about the disabilities. The class in itself group expression shows aggravated feelings about institutional and social disability harassment incidents. The limited e-participation of class in itself, persons with disabilities, shows that the accessibility metaphors are not adequate for the first level of e-participation (i.e., e-enabling) as the content accessibility is limited among the persons with disabilities. Those who availed of e-enabling conditions show their preparedness for the second level of e-participation (i.e., e-readiness) in attempts to bring awareness about the incidents. However, it is identified that this e-participation competency is not taking place at individual-self-reliant disabled persons, but mostly from organizations that fort for social inclusion. Hence, the class in itself competency is limited in nature.

The case of bullying Quaden, from Australia was shared on social media by the e-participants from across the globe. However, the series of disability discriminatory incident reporting during the health emergency is hardly viewed between 0 and 300 e-participants. The negative metaphor during a health emergency and the ableist compassionate expression toward the incidents can be merely a practice of capitalizing the disability incidents and lack of awareness of e-participants.

The study is limited to manual metaphorical analysis of awareness among e-participants to measure their e-participation competency. Study focus on health emergency pandemic, and future studies can integrate awareness during other forms of disaster. Only disability awareness was considered for analysis. However, since the disability inclusion improves diversity and innovations, this study is helpful for universal design principles for e-participation platforms. The current study emphasized native language since disability study demands regional understanding; future research will consider a large data set extracted automatically and provide more insight into the pandemic's e-participation competency.

**Acknowledgements** The scholar Dr. Sreejith Alathur is the awardee of ICSSR Special Call for Studies Focusing on Social Science Dimensions of Covid-19 Coronavirus Pandemic. This paper is largely an outcome of the Research project sponsored by the Indian Council of Social Science Research (ICSSR). However, the responsibility for the facts stated, opinions expressed, and the conclusions drawn is entirely that of the author.

## References

1. Spitzberg BH (1983) Communication competence as knowledge, skill, and impression. *Commun Educ* 32(3):323–329
2. Aral S, Eckles D (2019) Protecting elections from social media manipulation. *Science* 365(6456):858–861
3. Granovetter MS (1973) The strength of weak ties. *Am J Sociol* 78(6):1360–1380
4. Soldatic K (2013) The transnational sphere of justice: disability praxis and the politics of impairment. *Disabil Soc* 28(6):744–755
5. Mpofu E, DEd CRC, Oakland T (eds) (2009) *Rehabilitation and health assessment: applying ICF guidelines*. Springer Publishing Company
6. Mambrol N Disability studies, literary theory and criticism. Available at [literariness.org/2018/12/15/disability-studies-2/](http://literariness.org/2018/12/15/disability-studies-2/). Last accessed 23 April 2021
7. Schalk S (2013) Metaphorically speaking: ableist metaphors in feminist writing. *Disabil Stud Quart* 33(4)
8. Martin N (2012) Disability identity–disability pride. In: *Perspectives: policy and practice in higher education*, vol 16, no 1. pp 14–18
9. Friedner M (2008) Identity formation and transnational discourses (thinking beyond identity politics). *Indian J Gend Stud* 15(2):365–385
10. Shaw R, Kim YK, Hua J (2020) Governance, technology and citizen behavior in pandemic: lessons from COVID-19 in East Asia. *Prog Disaster Sci* 6(4):100090
11. OxCGRt Coronavirus government response tracker. Available at: [www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker](http://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker). Last accessed 25 April 2021
12. Kokoszka A (2007) *States of consciousness*. Springer, New York

13. Sommerhoff G (1989) Life, brain and consciousness, new perceptions through targeted systems analysis. In: *Advances in psychology*, vol 63. Elsevier Science & Technology, North-Holland
14. Zelazo PD, Moscovitch M, Thompson E (eds) (2007) *The Cambridge handbook of consciousness*, NY
15. Keefer LA, Landau MJ (2016) Metaphor and analogy in everyday problem solving. *WIREs* 7(6):394–405
16. Ellcessor E (2016) *Restricted access media, disability, and the politics of participation*, vol 6. NYU Press
17. Servick K For survivors of severe COVID-19, beating the virus is just the beginning, science. [www.sciencemag.org/news/2020/04/survivors-severe-covid-19-beating-virus-just-beginning](http://www.sciencemag.org/news/2020/04/survivors-severe-covid-19-beating-virus-just-beginning). Last accessed 09 April 2021
18. MSJE Exemption of employees with disabilities from essential services duty-reg, office memorandum. <https://dpe.gov.in/>. Last accessed 28 April 2021
19. Salloum A, Kondrat DC, Johnco C, Olson KR (2015) The role of self-care on compassion satisfaction, burnout and secondary trauma among child welfare workers. *Child Youth Serv Rev* 49(2):54–61
20. Olson KK (2005) Cyberspace as place and the limits of metaphor. *Convergence* 11(1):10–18
21. Harmon C, Kasa-Hendrickson C, Neal LVI (2009) Promoting cultural competencies for teachers of students with significant disabilities. *Res Pract Persons Severe Disabil* 34(3):137–144

# Fuzzy Logic Based Support Vector Regression (SVR) Model for Software Cost Estimation Using Machine Learning



Kamal Upreti, Umesh Kumar Singh, Rituraj Jain, Karamjeet Kaur,  
and A. K. Sharma

**Abstract** In this paper, a fuzzy logic (FL) based support vector regression (SVR) model along with COCOMO-II model by simulating machine learning is proposed and introduced. By using the benefits of fuzzy logic like adaption, accuracy in estimation, understandability and so on, can increase the estimation accuracy. Imposing fuzzy logic to any approach could lead to best performance by minimizing inaccuracy of input–output parameters. This proposed model uses fuzzy logic for overcoming the inaccuracy and ambiguity of parameters and provides high-accuracy results. More often, machine learning techniques such as ant colony optimization (ACO), K-Modes, Random forest and basic COCOMO model have been implemented and also comparison conducted in our proposed FL-COCOMO-II model. It shows that our proposed FL-COCOMO-II model provides better performance over the existing models. Therefore, the obtained results offer better accuracy, recall, execution time and runtime of the proposed model over the other existing techniques.

**Keywords** Software cost estimation · Machine learning · Support vector regression · K-modes · Random forest · COCOMO-II · Mean magnitude of relative error (MMRE) · Deliverable source instruction (DSI) metric

---

K. Upreti (✉)

Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology & Management, Delhi, India

e-mail: [kamal.upreti@adgitmdelhi.ac.in](mailto:kamal.upreti@adgitmdelhi.ac.in)

U. K. Singh

Institute of Computer Science, Vikram University, Ujjain, India

R. Jain

Department of Electrical and Computer Engineering, Wollega University, Nekemte, Ethiopia

K. Kaur

Department of Computer Science, MGM College of Engineering and Technology, Noida, India

A. K. Sharma

Department of CSI, University of Kota, Kota, India



# 1 Introduction

One of the most critical impressions on software quality collateral is that defects in software cost estimation would be a significant part of software development. In recent years, software cost estimation for assessing required effort has become a major problem. Cost estimation is a method or a figure of the achievable cost of a thing, program or an undertaking, enrolled in light of available information. Machine learning is advancing due to the algorithms it uses to make predictions. A number of algorithms are useful in forecasting and prediction. In calculating effort using any of the methods mentioned here, we require large data sets with maximum project information. What we need now is quick and accurate calculation, even if we have less. This paper uses machine learning to predict the effort if some parameters are known. The linear correlation and decision-making tree determine which parameters can help predicting. In software development lifecycle, accurate estimation of cost is one of the major aspects for every kind of project. This process of estimation involves predicting two things namely, the effort invested and the time taken for development. It is also a complicated task, as it supports the IT sectors to manage their software development. So far, a large number of models used for estimating cost that are employed in software developing organizations.

The main goal of estimating software cost is to maximize the probability rate of project success and to determine systematic efforts that are involved in software development life cycle. There are various techniques which have been used for optimization such as particle swarm optimization, ant colony optimization, differential evolution, bee colony optimization techniques, human opinion dynamics and COCOMO model, etc. Algorithms that are employed in software cost estimation are Firefly algorithm, BAT algorithm, Bayesian optimization algorithm, etc. [1]. VA standard metric to estimate cost of the software is introduced and followed. As these software cost metrics are in the form of measurements, quantization of certain things is involved. Mostly, the software estimation approaches require quantitative information to feed into the model such as size, productivity and various factors that are impacted in development of software. Software cost estimation is to predict the cost of resources originate for software development process.

Software cost estimation process is to originate the cost estimate of software. In software cost estimation, some factors are needed to estimate the cost are source line of code, scale drivers, cost drivers, analyst capability, programmer capability, software reliability, database size and execution time constraint. Software program is fed into the proposed model and the input SLOC is derived. It quantifies the software program's size. It is usually done by making source program lines counted. When counting library, comments, blank lines should be avoided. This metric can be applicable only in procedural language. It also counts defects, errors, POD, cost and deliverable source instruction (DSI) metric is same as SLOC. The only difference is 'if-else' statement. SLOC count that statement as one, but in DSI, it counted based on the statement iteration. From those inputs, along with cost and scaling factors, fuzzy crisp inputs are calculated. These crisp inputs are then fed into COCOMO

II with fuzzy logic. The three main steps which are applied in fuzzy logic system are fuzzification, rule-based system, and defuzzification. Fuzzification is a method of changing a crispy input to a fuzzy set. Rule-based system is a fuzzy system uses IF–THEN rules in their logic systems and defuzzification is a method of changing the output from fuzzy system into crisp output. Some optimization techniques are used in software cost estimation. In the first concept, basic COCOMO model is described using their estimation cost is described. In the second stage of this work, COCOMO II is described [2]. At the final stage of this research work, FL COCCOMO-II model is described and compared with other existing models such as ACO, K-modes, RF model and basic COCOCOMO model is presented. Thus, the outcomes show the accuracy, recall, execution time and runtime of the proposed framework over the other existing methods. Efficient software cost estimation is done using recall, accuracy, runtime and execution time comparison with existing models [3]. Here, the dataset used is FL-COCOMO-II where the execution time and runtime of proposed framework are low, and the accuracy is highly predicted.

The remaining content of the paper is given in the following: Sect. 2 introduces related literature review of the relevant field. Section 3 shows the proposed methodology along with the performance metrics, Sect. 4 covers findings and result analysis, and Sect. 5 contains conclusion and future scope with a list of references at the end.

## 2 Literature Review

Numerous researchers in the field of software cost estimation have carried out many researches using by using machine learning strategies. Some of these points are discussed further below.

In 2017, Tripathi et al. [4] published a paper demonstrating the importance of comparing traditional techniques to machine learning (ML) frameworks. Gained results show that machine learning approaches provide better quantification than conventional strategies. The creators have also described machine learning techniques and evaluation techniques for algorithm estimation.

In 2017, Jayaraj et al. [5] showed how difficult it is to estimate various costs with a complex action. Organizational assessments built on test methodologies and methodologies were always observed to incorporate new information.

In 2017, Akmal et al. [6] shows previous machine learning use enhances our programming abilities. Our goal was to look at previous studies in order to see if machine learning applied to information indexes such as basing them on three topics, whether information indexing philosophies such as using machine learning-based approaches were employed, and if these clustered instruments were critical to their success.

In 2016, Kumar et al. [7] shown a discussion about the machine learning computations, Naive Bayes, for instance, appear in connection to the use of software cost estimating, COCOMO II costs.

In 2016, Kibi et al. [8] suggested a merger in order to enhance COCOMO estimation performance. This technique has a significantly improved co-op COCOMO coefficient in comparison with the fundamental method of COCOMO calculation. Selecting the most effective coefficients multiplies the technique's effectiveness. Recreation shows up when MMRE and PRED (0.15). The experts obtained another type of COCOMO model as shown by streamlining the characteristics of the preceding exhibit with the addition of the specified number of attributes.

In 2015, Shrivastava et al. [9] published a paper that covers all of the software cost estimation models and techniques in use today. Also, the following features and weaknesses are currently seen in this field: Furthermore, it presents new procedures as well as their advantages and shortcomings. This paper describes a survey of various software cost estimation techniques. An adequate appraisal requires detailed and dependable information on all systems, and their basic programming information must be readily available.

In 2015, Bhatia et al. [10] presented their paper and discussed the M5P decision tree technique for evaluating effort in the field of software headway. With a focused on her objective of getting the job done, she concentrated on M5PT decision tree design. Improvements in decision tree characteristics can allow the effort screening system to be refined and implemented in a variety of ways.

In 2014, Sharma et al. [11] came up with an audit to check on different information tunneling approaches that had been employed. Various ML techniques (K-Means, K-NN, K-Nearest Neighbors, SVR), regression (MARS), and depiction techniques (CART) (SVM-Support vector machine, CBR-Case based thinking). These techniques can be used to enhance the calculation of exertion. Some data mining algorithms have been revealed to improve software estimation.

### **3 Proposed Methodology**

The performance metrics of our proposed fuzzy logic based SVR model for software cost estimation model are used as—recall, accuracy, runtime and execution time.

#### **3.1 Performance Metrics**

These parameters are compared with the existing models. FL-COCOMO II dataset have been implemented in this section.

##### **3.1.1 Accuracy**

Accuracy evaluates to the near measured value to a known or standard value.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \times 100$$

### 3.1.2 Recall

Recall is document percentage of relevant query that are retrieved. It is also known as completeness. Precision and recall are inversely related. Perfect score of recall is 1.0.

Recall is measured using the given formula:

$$\text{Recall} = \frac{TP}{(TP + FN)}$$

Accuracy of classifier is also evaluated based on the precision and recall values. TP-True Prediction is the positive part tuples that were noted accurately by the classifier, whereas TN represents the negative parts that were noted accurately by the classifier. FN-False Prediction negative is the negative parts that were noted wrongly as positive, whereas FN represents the positive parts that were noted falsely as negative [12].

### 3.1.3 Runtime

The total period of time the process is executing. The running time should also be lesser during the process both run time and the execution time are estimated in the seconds [3]. The run time is an interval between the beginning of the process and end of the process. Here, lesser time is a greater efficiency.

### 3.1.4 Execution Time

The execution time is defined as the method of manipulating the reliable service selection time with the more strength in security. Execution time is the time spent to finish the particular process and if comparison done the time should be lesser than the other process then it is measured that the execution time is good as well [13]. The execution time always predicted as lower value to find speed effort of software cost estimation and predicted execution time value in the form of process speed.

4 Findings and Result Analysis

In our proposed model, cost estimation has been conducted using recall, accuracy, runtime and execution time comparison with existing models. Here, the dataset used is FL-COCOMO II where the execution time and runtime of proposed system are low, and the accuracy is highly predicted. This chapter discusses the performance and simulation results of the prevailing and proposed system by estimating the result parameters [14]. The cost driver is the unit of activity that causes variation in activity cost. The values for some of the cost drivers of COCOMO-II model are shown in Table 1. It shows the list of some cost drivers and their range. There are several cost drivers but some of cost drivers and their particular range are mentioned. The above-mentioned drivers are FUSE, PLEX, TIME, TOOLS and SCED and examined their range in COCOMO II.

4.1 Datasets Selection

4.1.1 Dataset-1

From the economic opinion, the COCOCMO I dataset is a software engineering model which establish a cost estimation model from a dataset. The dataset 1 included sixty-three historical project data.

4.1.2 Dataset-2

From NASA Manger measurement dataset, the NASA93 data set contains ninety-three project data.

Table 1 Cost drivers of COCOMO-II model

Cost drivers	Range
Dataset size	0.90–1.28
Developed for reusability (FUSE)	0.95–1.24
Platform experience (PLEX)	1.19–0.85
Execution time limit (TIME)	1.00–1.63
Software tools (TOOLS)	1.17–0.78
Development schedule (SCED)	1.43–1.00

**Table 2** Comparative analysis of COCOMO II and FL-COCOMO-II model using MRE and PRED

Data sets	Models	Evaluation	
		Magnitude relative error (MRE)	Prediction (PRED) %
Dataset 1	COCOMO II	0.366149358	37
	FL-COCOMO II	0.393812753	34
Dataset 2	COCOMO II	0.341561862	33
	FL-COCOMO II	0.280274946	38
Dataset 3	COCOMO II	0.38219836	29
	FL-COCOMO II	0.32824384	33
Mean	COCOMO II	0.363303193	33
	FL-COCOMO II	0.334110513	35

### 4.1.3 Dataset-3

Four data sets form two corporate companies A and B. The project data is gathered from corporate companies A and B through a direct data collection in the COCOMO dataset format. Therefore, in this system, three different datasets of hundred and sixty projects are used for estimation of the FL-COCOMO II [11].

The FL-COCOMO II model is evaluated using three datasets. In order to calculate the accuracy for each dataset, the mean magnitude of relative error (MMRE) and PRED (25%) is used. The comparative analysis of the proposed FL-COCOMO-II and COCOMO model using three datasets are shown in Table 2.

The values of MREs and PREDs% for dataset 1 are calculated for both models mentioned above. In the above table, the dataset 1 shows that MRE of the FL-COCOMO II has lesser value and PRED of the same model has higher value. In this model, the COCOMO II originated higher accuracy for dataset 1 than FL-COCOMO II model.

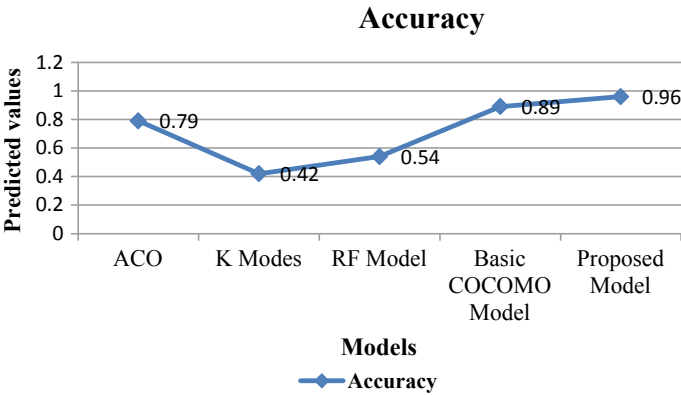
The values of MREs and PREDs % for dataset 2 are calculated for both models mentioned above. The dataset 2 shows that the MRE of FL-COCOMO II model has lesser value and PRED of the same model has higher value. In this proposed system, the FL-COCOMO II offers higher accuracy for dataset 2 than the COCOMO II model.

The values of MREs and PREDs% for dataset 3 are calculated for both models mentioned above. The dataset 3 shows that MRE of the FL-COCOMO has lesser value of MRE and higher value of PRED (%) than the COCOMO II model.

Finally, the values of MMREs of COCOMO II and FL-COCOMO II based on three data sets are 0.36 and 0.33, and the values of PRED (%) are 33 and 35 are evaluated. The obtained results show that the software cost estimation of MRE with lower value is an accurate estimate than a model with higher value and also the software cost estimation of PRED (%) with higher value is an accurate estimate than a model with lower value [14]. The estimation of both models is calculated based on the difference value. Table 3 shows the comparison of the performance metrics based upon estimation accuracy in COCOMO II and FL-COCOMO II model.

**Table 3** Comparison of performance metrics of FL-COCOMO-II and COCOMO-II

Models	Evaluation	
	Mean magnitude of relative error (MMRE)	Prediction (PRED %)
FL-COCOMO II	0.335110513	35
COCOMO II	0.36430193	33
Performance Metrics%	8.03	6.06



**Fig. 1** Accuracy comparison of proposed and existing models

In the above Table 3, the accuracy of the FL-COCOMO II model based on MMRE is 8.03% and PRED (%) is 6.06%. It is clearly showed that using FL-COCOMO II model, the accuracy is improved and it can be used alternatively for other software cost estimation model for accuracy estimation. Now, following Fig. 1 shows the graphical presentation of accuracy comparison of different existing models.

Figure 1 shows that the accuracy comparison of proposed model with existing models. Here, it is clearly proved that proposed model FL-COCOMO-II model has a higher value than other existing models such as ACO, K-Modes, RF model, COCOMO and their values are 0.79, 0.42, 0.54 and 0.89. The value of accuracy of the proposed model is 0.96 which clearly says the state of accuracy than other models. The higher value of accuracy is predicted in proposed FL-COCOMO-II model which is much efficient than the existing models [14]. Similarly, we have evaluated the performance metrics recall, runtime and execution time of our proposed fuzzy logic based SVR model for software cost estimation, and these parameters are also compared with the existing models. Hence, we describe the overall performance of our proposed fuzzy logic based SVR model for software cost estimation and comparative study of all existing models as presented below. Following Table 4 summaries the comparative result analysis of our proposed FL-COCOMO-II model with the four existing models.

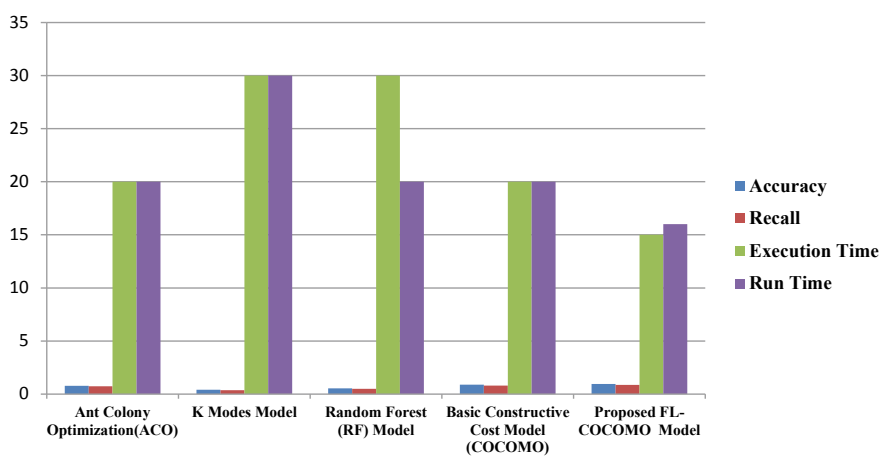
**Table 4** Result analysis of proposed FL-COCOMO-II model with existing models

Algorithm/Model used	Accuracy	Recall	Execution time	Run time
Ant colony optimization (ACO)	0.79	0.75	20	20
K-modes model	0.42	0.37	30	30
Random forest (RF) model	0.54	0.49	30	20
Basic constructive cost model (COCOMO)	0.89	0.81	20	20
Proposed FL-COCOMO model	0.96	0.87	15	16

Table 4 shows the overall comparisons of proposed as well as existing models. The models are estimated using accuracy, recall, execution time and run time. Table 4 summarizes the whole comparison with estimated values. Different models are referred and compared in this proposed model. This proposed FL-COCOMO-II model states that in every aspect it is higher than the other models. Accuracy, recall, execution time and run time are calculated using estimated software cost [1]. Time schedule and effort of software projects are also calculated in basic COCOMO model. The higher value of accuracy and recall, the lower value of execution time and run time is predicted as the efficient strength of proposed FL-COCOMO-II model [12].

In our proposed model, it is more efficient than the existing models such as ACO, K-Modes, random forest and basic COCOMO model. Here, the nearest value showed as basic COCOMO model and FL-COCOCMO-II model and comparison also conducted in our proposed FL-COCOMO-II model.

Thus, the graphical representation of comparative result analysis of our proposed FL-COCOMO-II model with the four existing models on four parameters such as—accuracy, recall, execution time and run time are shown in Fig. 2.



**Fig. 2** Comparison analysis of proposed FL-COCOMO-II model with existing models



## 5 Conclusion and Future Scope

Software cost estimation is concerned to develop an effort and to test the software that includes managing, training and buying extra equipment of software and servers. It is difficult and humans are terribly bad at predicting absolute outcomes. The prediction of cost of the resources is required to complete the software project work. Here, the quality problems, schedule overruns, outright cancelations and cost overruns are in very high frequencies that tend to large software projects. The customers and the developers are evaluated by the accurate of software cost. Scheduling, control, generating request for proposals, contract negotiations and monitoring are used in the software cost estimation. The under estimate of cost leads to bad quality, delay on time and ongoing functions. The formula for algorithmic technique was calculated. The cost factors are combined, and the models are developed. The proposed model FL-COCOMO II model is designed and implemented for software cost estimation to predict the better accuracy than other existing models. The fuzzy logic concepts are very high in the latest projects for software cost estimation. As a future scope, the fuzzy logic application in other software cost estimation models like differential evolution, human opinion dynamics, particle swarm optimization and BEE colony optimization algorithms may also be tested to explore software cost estimation.

## References

1. de Oliveira RAH, Stephan RM, Ferreira AC, Murta-Pina J (2020) Design and innovative test of a linear induction motor for urban MagLev vehicles. *Ind Appl IEEE Trans* 56(6):6949–6956
2. Mellis W, Basten D (2011) A current assessment of software development effort estimation. In: 2013 ACM/IEEE international symposium on empirical software engineering and measurement, Alberta, Canada, pp 235–244. <https://doi.org/10.1109/ESEM.2011.32>
3. Zhou X et al (2018) Cost estimation models of MJ class HTS superconducting magnetic energy storage magnets. *IEEE Trans Appl Supercond* 28(4):1–5
4. Tripathi R et al (2017) Machine learning methods of effort estimation and its performance evaluation criteria. *IJCSMC* 6(1):61–67
5. Jayaraj V et al (2017) Exploration of effort estimation techniques. *Int J Innovations Adv Comput Sci* 6(5):98–103
6. Akmal F et al (2017) A literature review study of software defect prediction using machine learning techniques. *IJERMT* 6(6):300–306
7. Kumar PM et al (2016) The effects of classification data mining algorithms on software cost estimation. *Int J Innovative Emerg Res Eng* 3(1):218–221
8. Khatibi BV et al (2016) An improved COCOMO based model to estimate the effort of software projects. *J Adv Comput Eng Technol* 2(2):11–22
9. Shrivastava A et al (2015) Software cost estimation: a survey of current practices. *Int J Eng Tech Res* 3(5):262–272
10. Bhatia S et al (2015) Implementing decision tree for software development effort estimation of software project. *Int J Innovative Res Comput Commun Eng* 3(5):4021–4025
11. Sharma M et al (2014) Software effort estimation with data mining techniques—a review. *Int J Eng Sci Res Technol* 3(3):1646–1653
12. Zhou X, Li S, Shi J, Zhang C, Tang Y, Zhang L, Gong K (2019) Capacity-control optimization of SMES in distribution networks with renewable energy. *Appl Supercond IEEE Trans* 29(5):1–6

13. Nawaz MA, Kumar M, Zaki MK, Gouda SK, Raj P, Kumar N, Das RB (2020) A methodology for software cost estimation using machine learning techniques. In: International conference on recent trends in artificial intelligence, IOT, smart cities and applications (ICAISC-2020)
14. Juneja N, Upreti K (2017) An introduction to few soft computing techniques to predict software quality. In: 2nd International conference on telecommunication and networks (TEL-NET), pp 1–6. <https://doi.org/10.1109/TEL-NET.2017.8343581>

# Author Index

## A

Abdul Rajak, A. R., 1, 11  
Abhishek, Manish Kumar, 225  
Abirami, A., 691  
Adi Narayana Reddy, K., 563  
Aditya, S., 303  
Afshar Alam, M., 21  
Agarkar, Samyak, 785  
Agarwal, Ayush, 55  
Agarwal, Parul, 167  
Aggarwal, Swati, 883  
Ahire, Pallavi, 197  
Ahmed, Jawed, 21, 167  
Ahmed, Rozha Kamal, 861  
Aishwarya, R., 65  
Aiswarya, V. B., 649  
Ajitha, M., 65  
Alagumeenaakshi, M., 639  
Alam, M. Afshar, 167  
Alathur, Sreejith, 907  
Alladi, Ramesh, 563  
Anekar, Shreya, 413  
Antoine, Gatera, 591  
Arora, Ritu, 761  
Arora, Shweta, 55  
Arun, C., 519

## B

Badgujar, Vishal S., 529  
Badjate, Mahima, 95  
Bag, Bikash Chandra, 481  
Bahurupi, Tejaswini, 797  
Bajpai, Gaurav, 591  
Balkothe, Soham, 113  
Balyan, Arnab, 883

Banerjee, Archita, 691  
Barik, Kousik, 691  
Bathija, Pranav, 345  
Bhamare, Devyani, 35  
Bhandari, Prashant, 773  
Bharathi, K. R., 575  
Bhat, Ashish, 345  
Bhattacharya, Samit, 737  
Bhope, Adwait, 773  
Bhuta, Nachiket, 611  
Bhuvaneshwari, M., 303  
Bilvashree, L., 575  
Bindu Bharati, G., 659  
Bisht, Narendra, 55  
Biswas, Saroj Kr., 703  
Bokhari, Hasnain, 459  
Borade, Jyoti G., 293  
Brinda Nadig, H. A., 575  
Buvanasri, A. K., 313

## C

Chakraborty, Imon, 751  
Chakraborty, Manomita, 703  
Chakravarthy, V. Deeban, 581  
Chandak, S. H., 357  
Chandra, Dara Sujith, 581  
Chandra Sekhar Reddy, V., 563  
Chandrashekar, Madan, 807  
Chatterjee, Soumyadip, 205, 217  
Chatuphale, Pallavi, 851  
Chawla, Harsh, 345  
Chhajed, R. R., 357  
Chheda, Dhairya, 247  
Choksi, Virti, 715  
Chugh, Nupur, 883

Colaco, John, 85

## D

Dalal, Het, 797  
 Dalal, Shweta, 715  
 Darji, Dhara N., 75  
 Das, L. N., 177  
 Das, Saptarshi, 691  
 Das, Soumen, 703  
 Dastoor, Sarosh, 715  
 Deepthi, S., 509  
 Deep, Yash, 807  
 Desai, Jaimin, 259  
 Deshpande, Arti, 345  
 Deshpande, A. V., 157  
 Deshpande, Prajakta, 95  
 Devika, S., 467  
 Dhamale, Rohan, 773  
 Dharmadhikari, S. C., 357  
 Dharme, Roshani K., 131  
 Divya, P., 649  
 Dixit, Mrudul, 441  
 Dongardive, Jyotshna, 851  
 Draheim, Dirk, 861  
 Durafe, Asha, 421  
 Dwivedi, Rajendra Kumar, 473

## E

Eskandari, Fatemeh, 861

## F

Fernandes, Priya, 611

## G

Ganeshan, M., 375  
 Gawai, Sachi, 851  
 Gawande, Ujwalla, 113  
 Gehani, Sakshi, 123  
 George Amalarethnam, D. I., 271  
 George, Hanna Susan, 467  
 Ghatwai, Sumedh, 113  
 Ghosh, Ramyani, 327  
 Ghosh, Sandeep, 481  
 Goswami, Shubham, 629  
 Govindu, S., 619  
 Gowthami, S., 659  
 Guhilot, Hansraj, 895

## H

Hande, Kapil, 785

Haripriya, M., 303  
 Harshitha, K., 575  
 Hegde, Vinayak, 551  
 Hiremath, P. S., 725  
 Hiremath, Shivanand C., 385

## I

Ilavarasan, P. Vigneswara, 751  
 Ingle, Purva, 365

## J

Jaikumar, V., 639  
 Jain, Jaysree, 539  
 Jain, Rituraj, 917  
 Jariwala, Siddhi, 797  
 Jayavel, Kayalvizhi, 591  
 Jha, Ravi Shankar, 139  
 John, Nithin Prince, 467  
 Johnson, Greshma, 467  
 Jose, Anjal A., 467  
 Joshi, Kavita Ajay, 55  
 Joshi, Mohit Hiteshkumar, 751  
 Joshi, Pratyush, 113  
 Joshi, Prutha, 773

## K

Kadhiwala, Bintu, 797  
 Kale, Sumit, 197  
 Kalkundri, Praveen, 895  
 Kalshetti, Urmila, 773  
 Kalyan Chakravathy, N. S., 639  
 Kamat, Vinod, 773  
 Kamble, Rutuja, 197  
 Kapadia, Parth, 797  
 Karthika, S., 313  
 Kaur, Karamjeet, 917  
 Kaur, Shubhpreet, 397  
 Kaur, Tarandeep, 397, 405  
 Kavyashree, B. S., 575  
 Kebede, Abdalganiy, 737  
 Kharat, Amit, 281  
 Kiwelekar, Arvind W., 187, 293  
 Koley, Chaitali, 481  
 Kompella, Shivani Krishna, 1, 11  
 Konar, Karabi, 691  
 Kosamkar, Pranali, 365  
 Kranthi Kumar Reddy, Y., 619  
 Krishnan, K. Keerthi, 819  
 Krishna Rao, L. V., 659  
 Kulkarni, Manish Gawali Viraj, 103  
 Kulkarni, Viraj, 281

Kulkarni, Vrushali, 365  
 Kumar, Akhilesh, 451  
 Kumaran, Santhi, 873  
 Kumar, N. S., 327  
 Kumar, Sanjeev, 629  
 Kunwar, Harsha C., 131  
 Kuradusenge, Martin, 873

## L

Li, Ming, 491  
 Lohani, Rajesh B., 85  
 Lokwani, Rohit, 281

## M

Madaan, Harshit, 103  
 Madhupriya, J., 271  
 Madli, Shiddalingeshwar, 807  
 Mahalle, Parikshit N., 157  
 Mahamunkar, Geetanjali S., 187  
 Maity, Amit, 481  
 Maity, Hirak Kumar, 481  
 Mallapur, Jayashree D., 385  
 Mandal, Surajit, 217  
 Mandhare, V. V., 35  
 Manjunath, K. A., 807  
 Marathe, Manisha, 773  
 Meenakshi, R., 313  
 Meghana, V. S., 327  
 Mhetre, Nalini, 157  
 Mikeka, Chomora, 591  
 Minani, Jean Baptiste, 873  
 Modi, Vedant, 365  
 Monu, 177  
 More, Nilkamal, 611  
 Musunuri, Surya, 807  
 Mutalik, Rohit, 247

## N

Naik, Kshitij, 797  
 Naik, Sapan, 237  
 Nainwal, Deepa, 55  
 Nainwal, Girish, 55  
 Narawade, Vaibhav E., 529, 539  
 Narula, Preet Nain Kaur, 55  
 Nemade, Shreyas, 113  
 Netak, Laxman D., 187, 293

## O

Oka, Tanvi, 441

## P

Paliwal, Sushila, 21  
 Pallavi, M. S., 551  
 Palve, Yogesh H., 131  
 Panda, Siba, 123  
 Pandey, Amrees, 451  
 Pandi, Chiranjeevi, 563  
 Pant, Aniruddha, 103, 281  
 Parikh, Satyen M., 75  
 Parveen, Suraiya, 21  
 Parween, Sagupha, 337  
 Patankar, Prachi, 441  
 Patel, Hiral R., 75  
 Patel, Toral, 95  
 Patidar, Vinod, 421  
 Patil, Akshita, 611  
 Patil, Prakashgouda, 725  
 Patnaik, Tsarina, 629  
 Pavan, Seemakurthi Sri Sathya, 581  
 Pawar, Tanmay D., 837  
 Pazare, Yash, 197  
 Phanindra, P. H., 619  
 Ponkshe, Saniya, 95  
 Prasad, Sneha, 629  
 Pratap, Anju, 467  
 Praveenkumar, S. M., 725  
 Premkumar, Anitha, 649  
 Priyenka, S., 519  
 Purkayastha, Biswajit, 703

## R

RadhaKrishnan, S., 509  
 Rajan, Annie, 601  
 Rajesh, A., 375  
 Rajesh, Varshini, 1  
 Rajeswara Rao, D., 225  
 Raj, Vinay, 681  
 Ramdasi, Ashish P., 45  
 Ramdey, Komal, 459  
 Ramesh, Rasya, 327  
 Rana, Yash, 237  
 Ravichandra, Sadam, 681  
 Ravi, Kalkundri, 895  
 Rege, Shivani, 441  
 Rohida, Sneha, 95

## S

Sabiha Naazneen, Sk., 659  
 Sadique, Jasim, 649  
 Sagarika, L., 519  
 Sahoo, Priti Ranjan, 139  
 Sai Charan, N. V. V. N., 619

Sai Vamsi, S., 659  
 Saiyed, Saziabegum, 149  
 Sajja, Priti, 149  
 Salgaonkar, Ambuja, 601  
 Sam, Kesia, 467  
 Sarda, Ekta, 539  
 Sarunisha, R., 303  
 Sathyalakshmi, S., 45  
 Seal, Dibyendu, 671  
 Seal, Reshmita, 671  
 Seth, Jitesh, 281  
 Sevak, Mayur M., 837  
 Shahin, Mahtab, 861  
 Shah, Karvin, 715  
 Shaikh, Zeeshan, 247  
 Shanthi Priya, R., 509  
 Sharma, A. K., 917  
 Sharma, Ashok, 397  
 Sharma, Yashvardhan, 761  
 Sheryl Oliver, A., 65  
 Shilpa, 405  
 Shilpa, M., 551  
 Shirisha, R., 649  
 Shrikhande, Nishant, 113  
 Siddique, Bushra, 827  
 Singh, Amandeep, 681  
 Singh, Chanavi, 365  
 Singh, Naveen, 499  
 Singh, Sachchidanand, 499  
 Singh, Subham, 481  
 Singh, Umesh Kumar, 917  
 Singh, Vinay, 499  
 Smrutirekha, 139  
 Sneha, S., 519  
 Soman, K. P., 819  
 Sonune, Harsha, 413  
 Sri Mounika, T. V. N. S. R., 619  
 Srivastava, Rashmi, 473  
 Suban, Stevin Mathew, 467  
 Sufyan Beg, M. M., 827  
 Sumithabhashini, P., 563  
 Suri, Riya, 611  
 Surywanshi, Jaya R., 131

Swaminathan, J. N., 639

## T

Tatari, Mona, 167  
 Tatte, Satyaki, 629  
 Thakkar, Viraj, 237  
 Toradmalle, Dhanashree, 247  
 Tripathy, Aruna, 337  
 Tyagi, Divya, 761

## U

Umamaheswari, S., 639  
 Upreti, Kamal, 917

## V

Valko, Andras, 591  
 Varghese, Salves, 467  
 Varghese, San Jos, 467  
 Venkatesan, Arivalagan, 807  
 Verma, Ravinder Kumar, 751  
 Vignaraj Ananth, V., 303

## W

Wanchoo, Kashika, 259

## Y

Yadav, Krishna, 197  
 Yadav, Prakhar, 451  
 Yadav, R. S., 451  
 Yadav, Vandana, 451  
 Yeginwar, Samrudhi, 413  
 Yoshitha, S., 649

## Z

Zanpure, Aashay, 773  
 Zaveri, Keyur, 715  
 Zennaro, Marco, 873