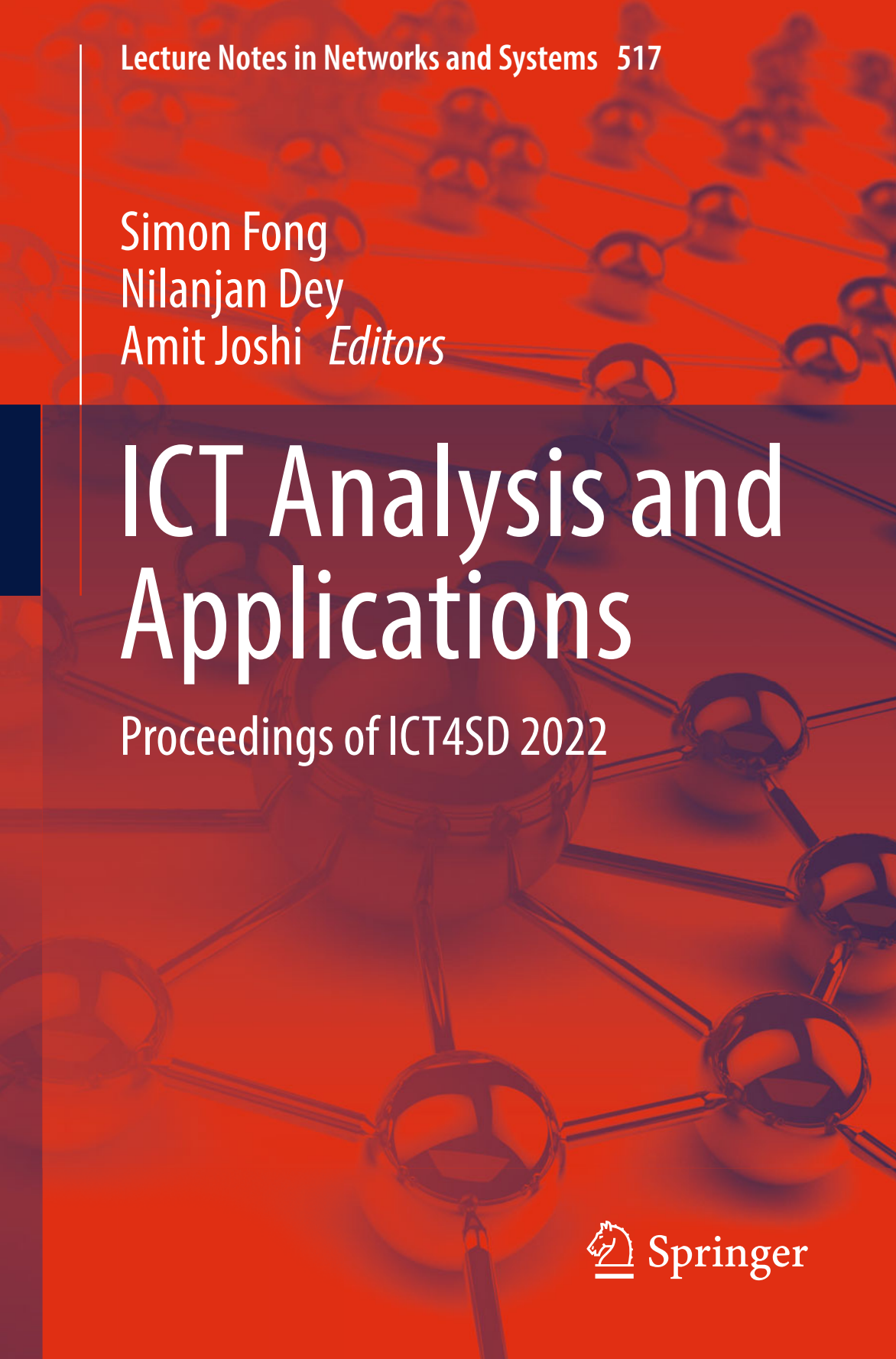
Lecture Notes in Networks and Systems 517

Simon Fong
Nilanjan Dey
Amit Joshi *Editors*

ICT Analysis and Applications

Proceedings of ICT4SD 2022



Springer

Lecture Notes in Networks and Systems

Volume 517

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.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Simon Fong · Nilanjan Dey · Amit Joshi
Editors

ICT Analysis and Applications

Proceedings of ICT4SD 2022

Springer

Editors
Simon Fong
University of Macau
Macau, Macao

Nilanjan Dey
JIS University
Kolkata, India

Amit Joshi
Global Knowledge Research Foundation
Ahmedabad, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-19-5223-4 ISBN 978-981-19-5224-1 (eBook)
https://doi.org/10.1007/978-981-19-5224-1

Preface

The Seventh International Conference on ICT for Sustainable Development (ICT4SD 2022) targets theory, development, applications, experiences and evaluation of interaction sciences with fellow students, researchers, and practitioners.

The conference is devoted to increasing 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 and provide guidance and address participant's questions and concerns.

The conference will be held on July 29–30, 2022, at Hotel Vivanta by TAJ, Panaji, Goa. The conference is organized by Global Knowledge Research Foundation, Supporting Partner InterYIT, International Federation for Information Processing, State Chamber Partner Goa Chamber of Commerce and Industry, and National Chamber Partner Knowledge Chamber of Commerce and Industry.

Research submissions in various advanced technology areas were received and after a rigorous peer review process with the help of program committee members and 193 external reviewers for 1200+ papers from 25 different countries including Algeria, United States, United Arab Emirates, Serbia, Qatar, Mauritius, Egypt, Saudi Arabia, Ethiopia, Oman, out of which 225 were accepted with an acceptance ratio of 0.19. These will be presented in 30 parallel sessions in two days organized physically and virtually including 1 inaugural and 1 keynote session.

Technology is the driving force of progress in this era of globalization. Information and communication technology (ICT) has become a functional requirement for the socio-economic 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 days conference had presentations from the researchers, scientists, academia, and students on the research work carried out by them in different sectors.

The conference has invited eminent dignitaries including Shri. Nitin Kunkolienker, President, MAIT, Prof. Mike Hinchey, President, IFIP, Dr. Milan Tuba, Vice Rector for International Relations, Singidunum University, Serbia, Prof. Jagdish Bansal, India, Mr. Aninda Bose, Springer Nature, and Dr. Amit Joshi, Director, G R Foundation.

Taipa, Macao — Simon Fong
Kolkata, India — Nilanjan Dey
Ahmedabad, India — Amit Joshi

Contents

Editors and Contributors

About the Editors

Simon Fong graduated from La Trobe University, Australia, with a 1st Class Honors B.Eng. Computer Systems degree and a Ph.D. Computer Science degree in 1993 and 1998, respectively. Simon is now working as Associate Professor at the Computer and Information Science Department of the University of Macau. Prior to his academic career, Simon took up various IT posts in Melbourne, Hong Kong, and Singapore. Dr. Fong has published over 396 international conference and journal papers, mostly in the areas of data mining and metaheuristics. He is also one of the founding members of the Data Analytics and Collaborative Computing Research Group in the Faculty of Science and Technology.

Nilanjan Dey is Associate Professor in Department of Computer Science and Engineering, JIS University, Kolkata, India. He is a visiting fellow of Department of Biomedical Engineering, University of Reading, UK, and Visiting Professor of Duy Tan University, Vietnam. He was an honorary Visiting Scientist at Global Biomedical Technologies Inc., CA, USA (2012–2015). He is Research Scientist of Laboratory of Applied Mathematical Modeling in Human Physiology, Territorial Organization of Scientific and Engineering Unions, Bulgaria. He is Associate Researcher of Laboratoire RIADI, University of Manouba, Tunisia. He is Scientific Member of—Politécnica of Porto. He is an affiliated member of Signal Processing and Computational Intelligence Applied to Power Systems (PSCOPE) research group, Federal University of Juiz de Fora, Brazil. He was awarded his Ph.D. from Jadavpur University in 2015. In addition, recently. he was awarded as one among the top 10 most published academics in the field of Computer Science in India (2015–17) during 'Faculty Research Awards' organized by Careers 360 at New Delhi, India. Before he joined JIS University, he was Assistant Professor of JIS College of Engineering and Bengal College of Engineering and Technology and Techno International New Town. He has authored/edited more than 85 books with Elsevier, Wiley, CRC Press and Springer and published more than 300 papers. His h-index is 51 with more

than 10,000 citations. He is Editor-in-Chief of *International Journal of Ambient Computing and Intelligence* (IJACI, IGI Global, UK, Scopus), Co-editor-in-Chief of *International Journal of Synthetic Emotions* (IGI Global, US, DBLP, ACM dl) and *International Journal of Natural Computing Research* (IGI Global, US, DBLP, ACM dl). He is Series Co-editor of Springer Tracts in Nature-Inspired Computing (STNIC), Springer, Series Co-editor of Advances in Ubiquitous Sensing Applications for Healthcare (AUSAH), Elsevier, Series Editor of Computational Intelligence in Engineering Problem Solving and Intelligent Signal processing and data analysis, CRC Press (FOCUS/Brief Series) and Advances in Geospatial Technologies (AGT) Book Series, (IGI Global), US, serves as an editorial board member of several international journals, including *International Journal of Image Mining* (IJIM), Inderscience, Associated Editor of *IEEE Access* (SCI-Indexed), and *International Journal of Information Technology*, Springer. His main research interests include medical imaging, machine learning, computer-aided diagnosis as well as data mining. He has been on program committees of over 100 international conferences, a workshop organizer of 15 workshops, and acted as a program co-chair and/or advisory chair of more than 15 international conferences. He has given more than 60 invited lectures in ten countries, including many invited plenary/keynote talks at the international conferences such as ITITS2017 (China), TIMEC2017 (Egypt) and BioCom2018 (UK), etc.

Dr. Amit Joshi is currently Director of Global Knowledge Research Foundation, also an entrepreneur and researcher who has completed his graduation (B.Tech.) in Information Technology and M.Tech. in Computer Science and Engineering and completed his research in the areas of cloud computing and cryptography in medical imaging with a focus on analysis of the current government strategies and world forums needs in different sectors on security purposes. He has an experience of around ten years in academic and industry in prestigious organizations. He is an active member of ACM, IEEE, CSI, AMIE, IACSIT-Singapore, IDES, ACEEE, NPA, and many other professional societies. Currently, he is Chairman of Computer Society of India (CSI) Udaipur Chapter and Secretary for Association of Computing Machinery (ACM) Udaipur Professional Chapter. Further currently, he is also 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. He has also edited more than 20 books which are published by Springer, ACM, and other reputed publishers. He has also organized more than 40 national and international conferences and workshops through ACM, Springer, and IEEE across five countries including India, UK, Thailand, and Europe. Apart from this academic involvement, he is also associated with the business and industry community across the globe. He is currently Director of Emanant TechMedia (P) Limited which focuses on ICT and web-based services. With Global Knowledge Research Foundation, his major work area is focused on building and making effective linkages among bureaucrats, industry associations, academic leaders and regulatory authorities for effective connectivity for taking up common issues related to research in different sectors and has organized more than 20

industry forum events and recognized and felicitated Awards Ceremonies through his organizations in last five years. Liaisoning/communications with state as well as central establishments along with corporate and academics to take industry and government sector together. He is majorly focusing on academic tie-ups for medical, engineering, and management sectors; for the same, he is representing the organization in many delegations to countries including USA, UK, Canada, and Israel and visited more than 15 countries. He always looks forward and strives for networking in various sectors with a focus on education and industry sector.

Contributors

Aarthi R. Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Achi Harrisson Thiziers Unité de Recherche et d'Expertise Numérique (UREN), Université Virtuelle de Côte d'Ivoire, Yamoussoukro, Côte d'Ivoire

Ahmed Rakib Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Alam Naser Abdullah University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Ali-Gombe Adamu Mintra Research, Aberdeen, UK

Amudha T. Department of Computer Applications, Bharathiar University, Coimbatore, India

Amutha C. Department of EEE, Rajalakshmi College of Engineering, Thandalam, Chennai, Tamil Nadu, India

Anand Vighnesh Department of Computer Sciences, College of Computer Science and Information Technology, King Faisal University, Al Hofuf, Saudi Arabia

Anni Nigar Sultana University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Ansari Mohmmed Rizwan Rajasthan Technical University, Kota, India

Apsangi Minal PVG's COET GKPIM, Pune, India

Arju Muhammad Ashfakur Rahman Daffodil International University, Dhaka, Bangladesh

Arulmozhi M. Department of ECE, Rajalakshmi College of Engineering, Thandalam, Chennai, Tamil Nadu, India

Auradkar Prafullata Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Aurpita Sumyea Nazifa University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Ayachit Isha MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

Bajaj Srinidhi Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

Banerjee Aishwarya Academy of Scientific and Innovative Research, Ghaziabad, India

Barpanda Nalini Kanta Department of Electronics, Sambalpur University, Burla, Odisha, India

Behera Santi Kumari Department of Computer Science and Engineering, VSSUT, Burla, India

Benti Dawit Milkiyas Department of Computer Science, Assosa University, Assosa, Ethiopia

Bhatia Gresha Vivekanand Education Society's Institute of Technology, Mumbai, India

Bhatt Harshita Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

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

Bhingarkar Sukhada Dr. Vishwanath Karad MIT—World Peace University, Pune, India

Bhuiyan Mazedul Haque BRAC University, Dhaka, Bangladesh

Bhutada Neil Dr. Vishwanath Karad MIT—World Peace University, Pune, India

Bidwai Sukrut Dr. Vishwanath Karad MIT—World Peace University, Pune, India

Biswas Ujjwal Indian Institute of Technology, Guwahati, Assam, India

Chaudhari Manisha Department of Computer Engineering, Indus University, Ahmedabad, India

Chaudhari Prajakta Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

Chaudhari Vrushali MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

Chaudhary Deevesh School of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India

Chebolu Abhishek Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

Choudhury Bidisha Department of Sanskri, Visva Bharati University, West Bengal, India

Chouhan Aryan Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Chugh Snehalraj Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

Cornet Souresh Amrita School for Sustainable Development, Amrita Vishwa Vidyapeetham, Amritapuri, India

Cunduri Mayra Alejandra Pacheco Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador

Dalwadi Neha Shri K. J. Polytechnic, Bharuch, Gujarat, India

Das Abhiroop Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Das Anupam Department of CSE, Royal Global University, Assam, India

Das Supratim Subhra Aliah University, Newtown Kolkata, West Bengal, India

Das Tanaya Adamas University, Kolkata, India

Deepika M. B. S. Abdur Rahman Crescent Institute of Science & Technology, Chennai, India

Deshmukh Neha Department of Information Technology, A.P. Shah Institute of Technology, Thane, Maharashtra, India

Desisa Eshetu Gusare Department of Computer Science, Assosa University, Assosa, Ethiopia

Deva Sarthak Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Devi Manju The Oxford College of Engineering, Bengaluru, India

Devikar Soniya MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

Dhaka Vijaypal Singh Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India

Dhakane Swapnali Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

Dhal Sunil Sri Sri University, Cuttack, India

Dhaneshwar Ritika University Institute of Engineering and Technology (UIET), Panjab University, Chandigarh, India

Dhanush K. Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Dhomane Mrunal G H Raisoni College of Engineering, Nagpur, India

Dhuri Soham Department of Information Technology, A.P. Shah Institute of Technology, Thane, MH, India

Diip Nazmuj Shakib University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Divya Udayan J. Amrita Vishwa Vidyapeetham, Amritapuri, Kerala, India

Dubey Aditya Maulana Azad National Institute of Technology, Bhopal, India

Dugad Shashikant Tata Institute of Fundamental Research, Mumbai, India

Dutt Suniti Chandigarh University, Punjab, India

Dutta Deep Dr. B. C. Roy Engineering College, Durgapur, India

Dwivedi Smrity Department of Electronics Engineering, IIT-BHU Varanasi, Varanasi, India

Ebenezer Abishek B. Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, India

Eedula Navya Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Estrella Diego Ramiro Ñacato Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador

Eswarawaka Rajesh Department of Artificial Intelligence and Machine Learning, AMC Engineering College, Bangalore, India

Fakir Al Imran Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Farahi Aishwariya Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Gadesha Vrunda School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, Gujarat, India

Gagan G. R. Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Gajjar Sachin Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Gandhi Kinjal Parul University, Waghodia, India

Ganesh C. Rajalakshmi Institute of Technology, Chennai, India

Ganesh K. Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Gogi Heena Kouser Department of ISE, Research Scholar PDACE, Kalaburgi, Karnataka, India

Gogul Nithish K. S. Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Gond Vikesh Kumar Maulana Azad National Institute of Technology, Bhopal, India

Gowthami B. V. Lakireddy Balireddy College of Engineering, Mylavaram, AP, India

Goyal Somya Manipal University Jaipur, Jaipur, Rajasthan, India

Guiguemde Jacques Rodrigue Equipe de Recherche Signal, Image et Communications (ERSIC) Laboratoire d'Algèbre de Mathématiques Discrétes et Informatique (LAMDI), Université Nazi BONI, Houet, Burkina Faso;
Réseau des Enseignants Chercheurs et Chercheurs en Informatique du Faso (RECIF), Ouagadougou, Burkina Faso

Gulati Archana School of Business Management, NMIMS University, Mumbai, Maharashtra, India

Harsha Deva Chandigarh University, Punjab, India

Hema B. Lakireddy Balireddy College of Engineering, Mylavaram, AP, India

Hiran Divya Department of Home Science, Govt. Meera Girl's College, Udaipur, India

Hiremath Tejas Dr. Vishwanath Karad MIT—World Peace University, Pune, India

Ingale Anupama K. Rajarambapu Institute of Technology, Isalampur, Maharashtra, India

Iyer Nandini G. Department of ECE, Rajalakshmi College of Engineering, Rajalakshmi Nagar, Thandalam, Chennai, Tamil Nadu, India

Jadhav Vanita D. SKNCOE, Savitribai Phule Pune University, Vadgaon(Bu), Pune, Maharashtra, India

Jagdale Devang Dr. Vishwanath Karad MIT—World Peace University, Pune, India

Jain Arpit A. GLS University, Gujarat, India

Jain Jinit Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Janani P. Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Janbhasha Shaik Department of Computer Science, Assosa University, Assosa, Ethiopia

Jaya T. VELS Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India

Jayalakshmi C. Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Jeevan Pradeep K. Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

Joshi Keyur School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, Gujarat, India

Kabore Kisito Equipe de Recherche Signal, Image et Communications (ERSIC) Laboratoire d'Algèbre de Mathématiques Discrétes et Informatique (LAMDI), Université Nazi BONI, Houet, Burkina Faso

Kadam Adwait PVG's COET GKPIM, Pune, India

Kailash M. Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Kalambur Subramaniam Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Kale Mohit Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India

Kallimani Jagadish S. Department of Artificial Intelligence and Machine Learning, AMC Engineering College, Bangalore, India;
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore and Affiliated to Visvesvaraya Technological University, Belagavi, India;
Department of Artificial Intelligence and Machine Learning, M S Ramaiah Institute of Technology, Bangalore, Karnataka, India

Kamalov Firuz Department of Electrical Engineering, Canadian University Dubai, Dubai, UAE

Karaka Teja C. Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Karmakar Pratap CSIR-Central Mechanical Engineering Research Institute, Durgapur, India

Karthika D. Patrician College of Arts & Science, Chennai, Tamil Nadu, India

Kaur Arshpreet Department of Computer Science & Engineering, Alliance College of Engineering and Design, Alliance University, Bengaluru, India

Kaur Mandeep University Institute of Engineering and Technology (UIET), Panjab University, Chandigarh, India

Kaur Manvjeet Cyber Security Research Centre (CSRC), Punjab Engineering College (PEC), Chandigarh, India

Kawatra Ruchi Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Keni Yash MIT Art, Design and Technology University, Maharashtra, Pune, India

Khare Nilay Maulana Azad National Institute of Technology, Bhopal, India

Kinger Shakti Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, Maharashtra, India

Kolhe Kishor School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Kone Tiemoman Unité de Recherche et d'Expertise Numérique (UREN), Université Virtuelle de Côte d'Ivoire, Yamoussoukro, Côte d'Ivoire

Kothari Hemant Pacific Academy of Higher Education and Research University, Udaipur, India

Kriplani Sejal Vivekanand Education Society's Institute of Technology, Mumbai, India

Krishna Rao L. V. Lakireddy Balireddy College of Engineering, Mylavaram, AP, India

Kumar Akshay Chandigarh University, Punjab, India

Kumar Ashish Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur, India

Kumar Dikshesh Chandigarh University, Punjab, India

Kurup Lakshmi Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Laha Arunava Dr B. C. Roy Engineering College, Durgapur, India

Lakshmi Narayana Hanumantha Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Luniya Diya Department of Information Technology, A.P. Shah Institute of Technology, Thane, MH, India

Maakar Sunil Kr. School of Computing Science and Engineering, Galgotias University, Greater Noida, India

Mahkalkar Namrata Priyadarshini College of Engineering, Nagpur, India

Makh Sanjay Sant Gadagebaba University, Amaravati, Maharashtra, India

Malagi Kiran B. Dayananda Sagar University, Bangalore, India

Mandal Souraneel Adamas University, Kolkata, India

Mante Jyoti School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India

Mayurappriyan P. S. Kumaraguru College of Technology, Coimbatore, India

Mazumder Chaitali Dr. B. C. Roy Engineering College, Durgapur, India

Mehra Jihan Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Mumbai, India

Memon Ahbaz Dr. Vishwanath Karad MIT World Peace University, Pune, Maharashtra, India

Mishra Deepanjali School of Humanities, KIIT University, Bhubaneswar, India

Mishra Pragnyaban Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

Mishra Wricha MIT Art, Design and Technology University, Maharashtra, Pune, India

Mohod S. W. Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

Momo Anika Tahsin University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Mounika Padire VELS Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India

Moussa Sherif Department of Electrical Engineering, Canadian University Dubai, Dubai, UAE

Mukherjee Arpita CSIR-Central Mechanical Engineering Research Institute, Durgapur, India

Mukherjee Debaprasad Dr. B. C. Roy Engineering College, Durgapur, India

Nachnani Dimple Vivekanand Education Society's Institute of Technology, Mumbai, India

Nagdive Ashlesha G H Raisoni College of Engineering, Nagpur, India

Naik Shefali School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, Gujarat, India

Naik Varsha Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, Maharashtra, India

Nair Manu Rajan Department of Computer Applications, Bharathiar University, Coimbatore, India

Nanda Deep Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Nandyal Suvarna Department of CSE, PDA College of Engineering, Kalaburgi, Karnataka, India

Narayanan Mukunth Sri Sri University, Cuttack, India

Narendra Babu G. Lakireddy Balireddy College of Engineering, Mylavaram, AP, India

Natarajan Mrudhulraj Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Nirmala Devi M. Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Padole Mamta The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat, India

Palai Golak Bihari School of Humanities, KIIT University, Bhubaneswar, India

Palwalia D. K. Rajasthan Technical University, Kota, India

Panchal Leena Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

Panwar Himanshi Pacific Academy of Higher Education and Research University, Udaipur, India

Parameswari R. Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Parihar Mahendra Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Mumbai, India

Patel Bhavin Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Patel Drashti Mahendrabhai Medusind Solutions, Ramanathapuram, India

Patel Hardik Mahendrabhai Vidush Somany Institute of Technology and Research, Kadi, India

Patel Meghal J. IBM, Ahmedabad, India

Patel Pinal J. Government Engineering College, Gandhinagar, India

Patil Abhinav MIT Art, Design and Technology University, Maharashtra, Pune, India

Patil Kiran MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

Patil Lalit V. SKNCOE, Savitribai Phule Pune University, Vadgaon(Bu), Pune, Maharashtra, India

Patil Saloni Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

Patil Savita P. Rajarambapu Institute of Technology, Isalampur, Maharashtra, India

Patni Aarohi Department of Computer Sciences, College of Computer Science and Information Technology, King Faisal University, Al Hofuf, Saudi Arabia

Pattni Kevin Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

Paul Bijan Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Pawar Mayur Vivekanand Education Society's Institute of Technology, Mumbai, India

Pethe Rahul G.H.Raisoni Institute of Engineering and Technology, Nagpur, India

Pol Rohan Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India

Prathap Kumar B. Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India

Presilla R. Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore and Affiliated to Visvesvaraya Technological University, Belagavi, India;
Department of Artificial Intelligence and Machine Learning, M S Ramaiah Institute of Technology, Bangalore, Karnataka, India

Priyatharishini M. Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Prodhan Md. Shiam University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Ragib Mufaddal Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, Maharashtra, India

Raina Shivani PVG's COET GKPIM, Pune, India

Rajeshwari S. B. Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore and Affiliated to Visvesvaraya Technological University, Belagavi, India;
Department of Artificial Intelligence and Machine Learning, M S Ramaiah Institute of Technology, Bangalore, Karnataka, India

Ranga Shashi V. Government Engineering College, Gandhinagar, India

Ranjith V. Dayananda Sagar University, Bangalore, India

Ransing Rasika Datta Meghe College of Engineering, Navi Mumbai, Maharashtra, India

Rasool Akhtar Maulana Azad National Institute of Technology, Bhopal, India

Ratha Ashoka Kumar Department of Electronics, Sambalpur University, Burla, Odisha, India

Ravichandran C. S. Sri Ramakrishna Engineering College, Coimbatore, India

Robbi Pratyusha Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Rodge P. R. Shivajirao S Jondhale College of Engineering, Mumbai, India

Saha Bajradeepon Dr B. C. Roy Engineering College, Durgapur, India

Saha Sajib Adamas University, Kolkata, India

Sai Saketh A. Lakireddy Balireddy College of Engineering, Mylavaram, AP, India

Saini Aditya A P Shah Institute of Technology, Thane(w), Maharashtra, India

Samridhi Ayushee Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Samsir Neha A. GLS University, Gujarat, India

Sanas Krutika Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India

Sankaranarayanan Suresh Department of Computer Sciences, College of Computer Science and Information Technology, King Faisal University, Al Hofuf, Saudi Arabia

Sant Kunal A P Shah Institute of Technology, Thane(w), Maharashtra, India

Santacruz-Sulca Fabricio Javier Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador

Saravanan R. Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Sarma Shikhar Kumar Department of IT, Gauhati University, Assam, India

Sathya R. Department of ECE, Rajalakshmi College of Engineering, Thandalam, Chennai, Tamil Nadu, India

Sawant Pratik PVG's COET GKPIM, Pune, India

Sekh Md. Asraful Aliah University, Newtown Kolkata, West Bengal, India

Sere Abdoulaye Equipe de Recherche Signal, Image et Communications (ERSIC) Laboratoire d'Algèbre de Mathématiques Discrétes et Informatique (LAMDI), Université Nazi BONI, Houet, Burkina Faso;
Réseau des Enseignants Chercheurs et Chercheurs en Informatique du Faso (RECIF), Ouagadougou, Burkina Faso

Sethy Prabira Kumar Department of Electronics, Sambalpur University, Burla, Odisha, India

Shaikh Afzal Anjuman-I-Islam's Kalsekar Technical Campus, Panvel, India

Shaikh Shaista Anjuman-I-Islam's A. R Kalsekar Polytechnic, Panvel, India

Shanbhag Ishaan Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, Maharashtra, India

Shanmugasundaram R. Sri Ramakrishna Engineering College, Coimbatore, India

Sharma Divya Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Shashvat Kumar Department of Computer Science & Engineering, Alliance College of Engineering and Design, Alliance University, Bengaluru, India

Shekhawat Deepika School of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India

Silva Danny Velasco Universidad Nacional de Chimborazo (UNACH), Riobamba, Ecuador

Singh Aastha Chandigarh University, Punjab, India

Singh Harmeet Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India

Singh Shivoham Pacific Academy of Higher Education and Research University, Udaipur, India

Singh Yash Chandigarh University, Punjab, India

Sinha Rishabh Sri Sri University, Cuttack, India

Sinwar Deepak Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India

Sitaram Dinkar Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India

Sodur Shanta Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India

Somani Ishan School of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India

Soni Ayush Maulana Azad National Institute of Technology, Bhopal, India

Sonune Harsha MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India

Sornali Jannatul Ferdouse Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Srivastava Devesh Kumar School of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India

Srivastava Satish Surendra Shivajirao S Jondhale College of Engineering, Mumbai, India

Sudharsan S. Department of ECE, Rajalakshmi College of Engineering, Thandalam, Chennai, Tamil Nadu, India

Surender R. Department of ECE, Rajalakshmi College of Engineering, Thandalam, Chennai, Tamil Nadu, India

Swain Sumeet A P Shah Institute of Technology, Thane(w), Maharashtra, India

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

Tamilselvi P. Sri Eshwar College of Engineering, Coimbatore, India

Tewari Abha Vivekanand Education Society's Institute of Technology, Mumbai, India

Tirumony Mir Nushrat Zahan University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Udeshi Nilay Department of Information Technology, A.P. Shah Institute of Technology, Thane, MH, India

Upadhyay Manisha Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Velonde Salonee Vivekanand Education Society's Institute of Technology, Mumbai, India

Vignesh O. Easwari Engineering College, Chennai, India

Vijay S. Department of Chemical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

Vyas Mahek Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Wakode Rashmi A. Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India

Zahangir Md. Afraim Bin University of Liberal Arts Bangladesh, Dhaka, Bangladesh

Survey of Data Processing Software Tools for Global Navigation Satellite System

Sachin Gajjar, Manisha Upadhyay, Mahek Vyas, Ayushee Samridhi, and Bhavin Patel

Abstract Satellite navigation system has come a way long from its initial stages when it was used for military applications to now in mobile devices worldwide. With the advancement in the satellite navigation system, various new navigation satellite constellations have been set up in space. Various data processing and analyzing tools have been developed for the systems. The number of this software, both online and offline, has been increasing due to which it has become essential to have a detailed comparative study on this software to design more efficient ones in the future. To this end, this paper surveys different software for global navigation satellite systems. The selection of the software for the survey is based on their attractiveness among scientists, results published in literature, and noteworthy characteristics and features. The survey work aims to assist scientists, researchers, and software developers in selection of an apt software for their work based on system requirements, supporting constellation, supported data format, price, size, strengths, and weaknesses. Software developers can further identify limitations of the existing software and overcome them.

Keywords Software tools for satellite navigation system · Bernese · gLAB · GNSS · GPS · OPUS · RTKLIB · Teqc

1 Introduction

Global Navigation Satellite System (GNSS) is a satellite navigation system for global coverage. It includes Global Positioning System (GPS) [1] from the USA, Global Navigation Satellite System (GLONASS) from Russia [2], Galileo [3] from Europe, BeiDou [4] from China, Quasi-Zenith [5] Satellite System (QZSS) from Japan, and Indian Regional Navigation Satellite System (IRNSS) [6] from India. It is a constellation of satellites that broadcast signals continuously with information of the time, orbit, and status. This information signal is processed at the user end with the help

S. Gajjar (✉) · M. Upadhyay · M. Vyas · A. Samridhi · B. Patel
Institute of Technology, Nirma University, Ahmedabad, Gujarat 382481, India
e-mail: sachin.gajjar@nirmauni.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_2

of software to find location and time information. A ground-based network helps in monitoring satellite orbit and time corrections to maintain accuracy in solutions. The GNSS technology market is ubiquitous. Applications range from simple handy navigation aids with meter-level positioning to robust, centimeter-level positioning solutions for survey, unmanned, military, marine, and aerial navigation. In satellite communication tropospheric errors, ionospheric errors, weather conditions, receiver components delay, clock drift, and clock bias are affecting the accuracy of the navigation system. Hence, software tools are required to process the received data to get accurate results and evaluate the positioning performance of any navigation system. The popular software tools available from different sources are RTKLIB, Bernese, gLAB, Teqc, Terrapins, and GPSTk. The main purpose of these software tools is (i) to convert received raw data from a receiver into RINEX format, (ii) to preprocess the data by applying necessary atmospheric, clock, antenna, clock bias, etc., corrections defined by models, (iii) to evaluate position solution and performance using different positioning modes like differential GNSS (DGNSS), real-time kinetic (RTK), satellite-based augmentation systems and precise point positioning (PPP), and (iv) to analyze results using graphs and plots. The rest of the paper is organized as follows: Sect. 2 discusses the desktop-based software and their features. Exhaustive comparison of the different software is presented in Sect. 3. The paper is concluded in Sect. 4.

2 Desktop-Based Software

A number of software packages that can process precise GPS/GNSS data for research and other applications have been developed by different international research groups. The constellation support, system requirements, amount of facilities in the software, the availability of source code, documentation, and user support vary with individual development group. Some of the desktop-based software are discussed next.

2.1 Bernese

Bernese is a commercial GNSS software developed at the Astronomical Institute of the University of Bern (AIUB). As it provides high-precision multi-GNSS data processing, it comes under high-quality geodetic software packages for the post-processing mode. Widely used in more than 700 institutes for processing the software package comprises about 1500 modules giving inherent flexibility [7]. The package can be run on any system independent of an operating system. In addition, a set of executables for the programs and a platform-independent GUI is programmed using C++ using Qt4 library is provided for Windows OS users, while it is to be manually compiled in Linux environment. As numbers of permanent GNSS tracking networks

were established for various purposes, it resulted in large amount of data collected per day, which demanded a highly automated GNSS data processing. This demand led to the development of a new front end for automization, i.e., Bernese processing engine (BPE) [7]. Written in Perl, BPE increased the capabilities of Bernese to a new level. BPE follows a server–client-based architecture, where server and client communicate using TCP/IP protocol. Server is designed using C++ while client is written in Perl. It fully supports both American GPS and Russian GLONASS and Galileo to some extent. Future planning is to extend it to other available navigation systems. Being a commercial software, license model is distinguished between two types of users: (i) researcher where mostly universities or other institutions engaged in research activities can get the software at a reduced price but with a restriction of not selling product directly or indirectly and (ii) commercial users involving companies or public institutions who are focused on commercial result rather than scientific research [8]. Salient features are as stated below:

- Support for Low Earth Orbital (LEO).
- Single task on several central processing units (CPUs) and also different task on different CPUs [7].
- Support for L1, L2, and L5 band of frequencies [7, 9].
- Integrated processing of data from GPS and GLONASS receiver is possible.
- Processing of data can either be done in double-difference mode (setting up the difference between a pair of receivers and pair of satellites) or in the zero-difference mode.
- Phase defects can be done using different ambiguity resolution up to several thousand kilometer-long baseline available for GPS and GLONASS [8].
- Ready-to-use examples of direct processing are available in package [7].

However, on the downside, Bernese is a licensed and a closed source software.

2.2 RTKLIB

RTKLIB is a free, open-source software developed by Tomoji Takasu of the Tokyo University of Marine Science and Technology [10]. RTKLIB was originally developed for GPS but now is capable of supporting precise positioning with GLONASS, Galileo, QZSS, BeiDou, SBAS, and also IRNSS in their latest version. The package includes the executables graphical user interface (GUI) and character user interface (CUI) application program [11, 12]. Windows OS can execute GUI and CUI both, while Linux and Mac OS can only compile CUI. CUI application programs (APs) are written in ANSI C and GUI APs use C++ along with BORLAND Visual component Library for GUI toolkits [11]. RTKLIB APs include:

- RTKNAVI: It provides real-time processing for GNSS receiver logged raw data.
- RTKPOST: It takes standard RINEX observation data and GNSS navigation message file as input and computes and provides positioning solutions by using numerous position modes.

- RTKPLOT: RTKPLOT is used for viewing and plotting the positioning solution given by RTKPOST, RTKNAVI, and also for RINEX observation data and satellite visibility analysis with the graphical user interface.
- STRSVR: It is a communication utility used for setting a communication link for the user to transfer and receive input and output data stream.
- RTKCONV: It can be used for the conversion of RTCM and BINEX message format to receiver independent exchange format (RINEX). Also, it can extract SBAS messages from the receiver raw data.
- SRCTBLBROWS: It is a browser that provides access to the NTRIP source table. Networked transport of RTCM via Internet protocol (NTRIP) is a communication protocol for interchanging GNSS logged data [11].

Salient features are as stated below:

- Support for standard formats and protocols of GNSS RTCm, BINEX, NTRIP 1.0, NMEA 0183, SP3, ANTEX, and IONEX [11].
- Supports positioning modes with GNSS for both real-time and post-processing like DGPS/DGNSS, kinematic, fixed, PPP kinematic, PPP static and fixed, etc.
- A portable program library and several APIs are available along with the package.
- Supports external communication like serial, TCP/IP, NTRIP, and local log file. It uses for automatic download FTP/HTTP [11, 12].

2.3 GLAB

The GNSS-Lab Tool suite (gLAB) is an educational multi-purpose package for processing and analyzing GNSS data. It was developed by the research group of Astronomy and Geomatics (gAGE) from the Universitat Politecnica de Catalunya [13]. The software tool was initiated with the help of the European Space Agency as a part of the ESA satellite navigation educational program (EDUNAV) and is freely available to the GNSS researchers and students by the ESA Education Office [14]. It aims to aid students/researchers by providing educational tools and materials that teach GNSS both in a theoretical and practical way. The tool is distributed as a package of software along with tutorials and data set. The software comprises a set of binary files able to read GPS RINEX data and process it to show results in the form of graphics and data files. The manual consists of GNSS Fundamentals and Software routines of GNSS processing concepts like data processing and analyzing the results. With the latest release, it has been expanded to SBAS, Galileo, and GLONASS system. gLAB can be run on Linux (Ubuntu 10.04 and above), Windows platform as well as on Mac Operating Systems with a minimum system requirement of 1 GB of RAM, CPU with at least 1 GHz and 1 GB of free space in the hard disk for installation [15]. The software tool is subdivided into three independent executable as follows:

- Data processing core (DPC): It is programmed using C (C99) language which implements all data processing algorithms to process GNSS data and can be executed in a command window. It is highly configurable as per the user's requirement.
- Data analysis tool (DAT): Python is used to program this module. It uses output files generated by DPC to generate plots for data analysis. Like DPC, it is also highly configurable, and the user is free to change or add scale, color, plotting condition, etc. It supports multiple image formats in exporting output plots [16].
- Graphic user interface (GUI): It is written with the help of object-oriented programing language C++ and Qt. It acts as an interface to manage between DPC and DAT modules. It executes the DPC and DAT files as per the user's choice of parameter [15].

Salient features are as stated below:

- Supports standard formats like RINEX-3.00, SP3, ANTEX, and SINEX files.
- Automatic detection of RINEX format [16].
- Jump in carries phase measurement known as cycle slip that can be detected which removes the ambiguity of integer number of cycles.
- It applies various error correction models like satellite time, atmospheric delay, etc., along with the standard position system [13].

On the downside of the software, for getting the output, it requires both observation and navigation data.

2.4 Teqc

Teqc is a freeware software developed and managed by the UNVACO organization. With its final release on February 25, 2019, Teqc has reached its end of life with no further release and update. However, email support is provided by UNVACO [17]. Teqc stands for translational, editing, and quality checking. It is capable of translation of native binary format files like RINEX file creation. Editing/correcting of RINEX header metadata can alter or add information, including time windowing. Also, Teqc QC (quality checking) provides a quality check of GPS/GLONASS data (BINEX or RINEX observation files). All the above three parts of Teqc can function separately, in pairs or all together in one go. Software is available in the form of executable binaries and can only be executed through the command line. Operating system supporting Teqc includes Mac OS X 10.5 (Leopard) or later, x86 Linux, 32 and 64-bit, Windows 32-bit, Windows 64-bit for 7 onwards, Raspberry Pi, Viper ARM, and Arietta G25 [18].

Salient features are as stated below:

- Support of standard file formats such as BINEX, RINEX, IGS RTIGS, u-blox UBX, Trimble RT17, RT27, Canadian Marconi Binary, ARGO, Rockwell zodiac binary, and many more [18].
- Quality check of RINEX files with or without navigation files.
- Cut, slice, or join for two or more RINEX files.
- Produces ASCII-based time plot for processed data.

On the downside, a non-graphical user interface and no plotting tool facility make it less user friendly.

3 Comparison

Table 1 on the previous page compares all the software discussed in the paper. For an user well versed with the subject, RTKLIB is most suitable for windows platform as it easy to use yet efficient with many functionalities, whereas for a novice user, gLAB is more suitable as its GUI provides a simple interface with a tooltip facility for all of its features. Collectively on the basis of all the parameters, RTKLIB is best suited for processing applications as it supports many formats, protocols, positioning algorithms, and modes both in real and post-processing with nearly all constellations. After that comes gLAB because it is a user-friendly interface, which makes it easy to operate for beginners. Web-based GNSS services are also being developed by several organizations. Online services are mostly based on two types of approaches: precise point positioning (PPP) and relative solution approach. With web-based software, user has to submit data (preferably in RINEX format) using either their webpages or to FTP sites via the Internet. The software does the analysis and sends the results to the user via email.

4 Conclusion

This paper surveys different software used for processing and analyzing data from global navigation satellite systems. These tools are also used for (i) precise point positioning (PPP) of the navigation system, (ii) preprocessing for applying atmospheric, clock, antenna, clock bias, etc., corrections, and (iii) post-processing for evaluating receiver position using different positioning modes. Some of the other GNSS processing software not discussed in this paper include GPSTK [19], GipsyX [20], GAMIT/GLOBK [21], and RTnet [22]. From the software compared in this paper, RTKLIB is a constantly evolving open-source software that is best suited for processing applications as it supports a number of file formats, protocols, positioning algorithms, and modes.

Table 1 Comparison of software tools for global navigation satellite system

Software	Bernese	RTKLIB	gLAB	TEQC
Latest version	Bernese 5.2	RTKLIB 2.4.3	gLAB v5.4.4	TEQC Finis
Year of release	2016/12/16	2019/08/19	2019/05/10	2019/02/25
Developer	Astronomical Institute, University of Bern	Tokyo University of Marine Science and Technology	Group of Astronomy and Geomatics, Spain	UNVACO, USA
URL	http://www.berneseu-nibe.ch/	http://www.rtklib.com/	https://gageupc.ed-u/gLAB/	https://www.unavco.org/software/data-processing/teqc/
GNSS support	GPS, GLONASS, Galileo	GPS, GLONASS, Galileo, QZSS, BeiDou, SBAS, IRNSS	GPS, SBAS, GLONASS, Galileo	GPS, GLONASS, Galileo, SBAS, Beidou, QZSS, IRNSS
Platforms	Windows, Linux, Mac, Unix	Windows, Linux, Mac, Unix	Windows, Linux, Mac OS	Linux, Windows, Mac
GUI/CUI Support	GUI and CUI	GUI and CUI	GUI and CUI	CUI
Licensing terms	Commercial	Free	Free	Free
Language of source code	C++, Qt4	C, C++	C, Python, C++, Qt2	C, C++
Processing modes	Pre and Post-processing	Pre and Post-processing	Pre and Post-processing	Pre and Post-processing
File format supported	RINEX with LEO extension, SP3c, SINEX, IONEX, Clock RINEX, Troposphere SINEX, ANTEX, IERS ERP	GNSS RTCm, BINEX, NTRIP 1.0, NMEA 0183, SP3, ANTEX, IONEX	RINEX-3.00, SP3, ANTEX, SINEX	u-blox UBX, BINEX, RINEX, IGS RTIGS, Trimble RT17, RT27, Septentrio Binary Format, Javad JPS, Topcon TPS, Canadian Marconi Binary, ARGO, Rockwell zodiac binary
PPP support	Yes	Yes	Yes	Yes
Strong points	User-friendly GUI, highly precise, inherent flexibility due to a highly modular design, built-in HTML-based help system	User-friendly GUI, supports real-time, supports major features available in a commercial software	User-friendly GUI, most suitable for beginners, full traceability of intermediate computations, additional tooltip feature in GUI to explore the tool	Features for checking and correction in RINEX data, most suitable for basic preprocessing data
Weak points	Commercial, closed source, does not support IRNSS, SBAS, BeiDou, QZSS	Too complex for beginners, less flexible in settings for the coordinate system (default WGS-84), at a time, data processing can be done for one rover and one base station only	Closed source, does not support IRNSS, BeiDou, SBAS	Closed source, GUI not available, data plotting is not supported

Acknowledgements This work is supported by Indian Space Research Organization, Government of India, under Respond Scheme contract No. ISRO/RES/3/870/20-21.

References

1. Moorefield FD (2020) Global positioning system standard positioning service performance standard 5th edn. Department of Defense, Washington, DC (1):1–196
2. Korolev: Global navigation satellite system glonass—open service performance standard (os ps). Information and Analysis Center for Positioning, Navigation and Timing, Russia Edition 2.2(1):1–56
3. Hecker P, Bestmann U, Schwithal A, Stanisak M (2018) Galileo satellite navigation system—space applications on earth. European Parliamentary Research Service, Rue Wiertz 60, B-1047 Brussels, Belgium 1(1):1–60
4. Beidou navigation satellite system signal in space interface control document (2020) China Satellite Navigation Office. Beijing, China 1(1):1–57
5. Quasi-zenith satellite system performance standard (2020) Cabinet Office. Government of Japan 1(1):1–33
6. Mruthyunjaya L (2017) Indian regional navigation satellite system—signal in space ICD for standard point positioning service. ISRO Satellite Centre, Indian Space Research Organization, Bangalore 1.1(1):1–72
7. Dach R, Lutz S, Walser P, Fridez P (2018) Bernese GNSS software manual. Astronomical Institute, University of Bern 5.2(1):1–17
8. Dach R, Lutz S, Walser P, Fridez P (2015) Bernese GNSS software version 5.2 user manual. Astronomical Institute, University of Bern 5.2(1):3–8
9. Dach R, Bock v, Jäggi A, Lutz S, Meindl M, Meyer U, Orliac E, Ostini L, Prange L, Schaer S, Sosnica K, DT, Walser P (2010) Support from the bernese software for the EU reef-reprocessing. 7th EUREF local analysis centres workshop, Military University of Technology, Warsaw, pp 1–24
10. Takasu T (2009) Rtklib: open source program package for GNSS positioning 1:1–24
11. Takasu T (2013) Rtklib ver. 2.4.2 manual. Tokyo University of Marine Science and Technology, Tokyo 1(1):1–75
12. Ruddick R, Choy S (2018) An introduction to RTKLIB open source GNESS processing software. HKMO Technical Seminar, Istanbul 1:1–30
13. Hernandez-Pajares M, Juan JM, Sanz J, Ramos-Bosch P, Rovira-Garcia A, Salazar D, Ventura-Traveset J, Lopez-Echazarreta C, Hein G (2010) The ESA/UPC GNSS-lab tool (GLAB): an advanced multipurpose package for gnss data processing. In: 5th ESA workshop on satellite navigation technologies and European workshop on GNSS signals and signal processing, pp 1–8
14. Ibáñez D, Rovira-Garcia A, Sanz J, Juan J, González-Casado G, Jimenez-Baños D, López-Echazarreta C, Lapin I (2018) The gnss laboratory tool suite (glab) updates: Sbas, dgnss and global monitoring system. In: 9th ESA workshop on satellite navigation technologies and European workshop on GNSS signals and signal processing, pp 1–11
15. Ramos-Bosch P, Rovira-Garcia A, Ibáñez D (2020) GNSS-lab tool software user manual. ESA GNSS Education, Research group of Astronomy and Geomatics, Barcelona Version 1(15):1–70
16. Dionisio C, Citterico D, Pirazzi G, Quattro ND, Cucchi L, Marracci R, Valdambrini N, Formaioni I (2010) glab a fully software tool to generate, process and analyze gnss signals. In: 5th ESA workshop on satellite navigation technologies and European workshop on GNSS signals and signal processing, pp 1–7
17. Estey LH, Meertens CM (1999) TEQC: the multi-purpose toolkit for GPS/GLONASS data. GPS Solutions 3(1):42–49

18. Estey L, Wier S (2014) TEQC tutorial: basics of TEQC use and TEQC products. UNAVCO, Boulder, Colorado pp 1–61
19. Harris RB, Mach RG (2007) The GPSTK: an open source GPS toolkit. GPS Solutions 11(2):145–150
20. Bertiger W, Bar-Sever Y, Dorsey A, Haines B, Harvey N, Hemberger D, Heflin M, Lu W, Miller M (2020) GipsyX/RTGx, a new tool set for space geodetic operations and research. Adv Space Res 66(3):469–489
21. Herring TA, King RW, Floyd MA, McClusky SC (2018) Introduction to gamit/globk. Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology Release 10.7:1–9
22. GPS Solutions Inc. Boulder C (2021) Rtnet (real time network processing engine) software brochure

POS Tagging for the Primitive Languages of the World and Introducing a New Set of Universal POS Tagging for Sanskrit

Anupam Das, Bidisha Choudhury, and Shikhar Kumar Sarma

Abstract The digital structuring of a language depends on how their parts of speech are classified to make the language the most compatible for the user. The parts of speech (POS) of a language is defined as the word classifiers which classify a word more precisely for using it in the sentence properly for expressing the emotions or feelings of human beings through natural languages. The natural language processing in the research arena already classifies the exhaustive forms of POS of many languages in the world. Among all such languages, the POS of SANSKRIT language is also introduced by many eminent scholars. Here, an attempt is made to give the new set of POS of the SANSKRIT language with their proper definitions and explanations with examples and they are termed as Universal POS (UPOS). The set so defined that universally everybody can understand, and the further use of these tags can be made available to the researchers of the world. The new set is highly expected to be the most exhaustive form.

Keywords UPOS-universal POS · Primitive language · Word order · Compound

1 Introduction

In ancient times, human beings communicated their emotions and feeling or their instructions through gestures and some primitive oral sounds. Though it is difficult to give the exact period of time when the structured language came to the human society but linguists claimed around 10,000 years ago; the existence of the structured language evidenced through scriptures. The first structured language is really debatable as different linguists claimed differently. During the studies of linguists,

A. Das (✉)
Department of CSE, Royal Global University, Assam, India
e-mail: adas_arya@rediffmail.com

B. Choudhury
Department of Sanskri, Visva Bharati University, West Bengal, India

S. K. Sarma
Department of IT, Gauhati University, Assam, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_3

one thing came out as evidential that structured languages are used in their first appearances in texts formed and its use as formal verbal forms in contemporary era. Table 1 gives the five oldest languages claimed by many linguists without any further controversies [1].

The parts of speech (POS) of any language are the detailed word classifier. The best set of POS of a language can be claimed only when the POS covers the best possible sentence composition with its fullest meanings. The smallest entities used in each sentence are distinctly defined as the POS with their specific usages [1]. Table 2 gives the number of POS of the top 10 most popular languages of the world:

In the table, it is observed that most of the languages have their POS in very limited numbers with standard Chinese has 13 as maximum and Arabic has 3 as minimum. All these languages are structured by the grammarians with the proper

Table 1 Primitive languages and its existence

Ranks according to the evidence of scripts	Language name	Period of existence claimed	Evidence got as scripts	Development of parts of speech tags (POST)
1	Sanskrit	2nd millennium BC	2000 BC	2023 (including sub-tags)* introduced in this study
2	Greek	2000 BC	1500 BC	165 (including sub-tags)
3	Chinese	1500 BC	1250 BC	34 (including sub-tags)
4	Hebrew	1200 BC	1000 BC	23 (including sub-tags)
5	Tamil	2500 BC	300 BC	71 (including sub-tags)

Table 2 Primitive languages and their number of POS tags and number of SPEAKERS

S No.	Language name	No. of POS	Speakers
1	Standard Chinese	13 = 10 content words + 3 function words	1.1 billion
2	English	8	983 million
3	Hindustani	8	544 million
4	Spanish	9	527 million
5	Arabic	3	422 million
6	Malay	4	281million
7	Russian	9 = 4 main + 5 minor	267 million
8	Bengali	8	261 million
9	Portuguese	10	229 million
10	French	8	229 million

definitions of POS. The scopes of new classification of words in these languages are closed, and thus, they can be categorized as the closed languages. A language can be termed as open language, where the scope of inserting new set of POS is opened. The SANSKRIT falls under the open category language. There are many researchers tried to insert new set of POS into this language and successfully did it. In this paper, we tried another attempt to introduce a new set of POS called UPOS in the most structured language SANSKRIT. Here, we are not only inserting new set of POS but also digitizing all the POS with their respective tagging.

1.1 New Set of POS in Sanskrit

The POS in Sanskrit is crafted by many grammarians and linguists. There are certain issues which are not dealt with those works: like sub-tagging in main tags is not classified, and also sub-tag like case and case-ending was $b =$ not discussed. There are 100 main tags and total 2023 main tags along with their sub-tags are classified in this work.

1.2 Analysis of UPOS in Sanskrit

According to the sub-features of all the above tags are calculated, then the total number of UPOS will be as shown (Figs. 1, 2, 3, 4, 5, 6, 7).

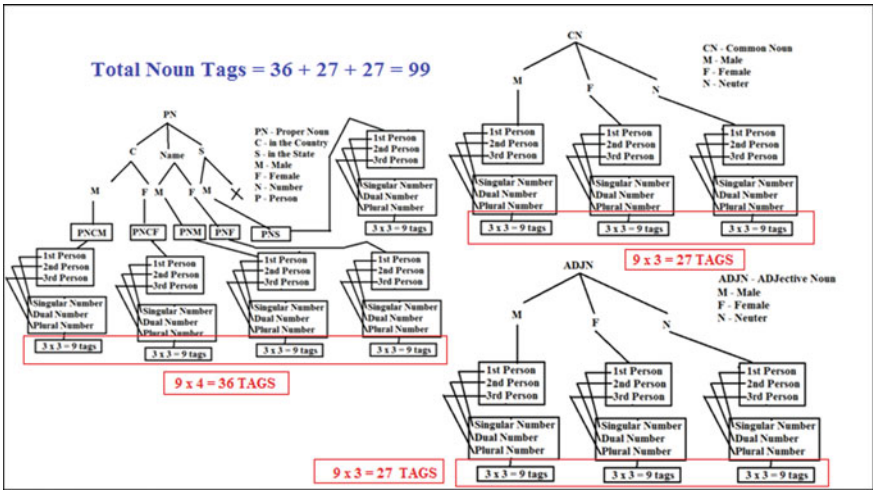


Fig. 1 Noun tags

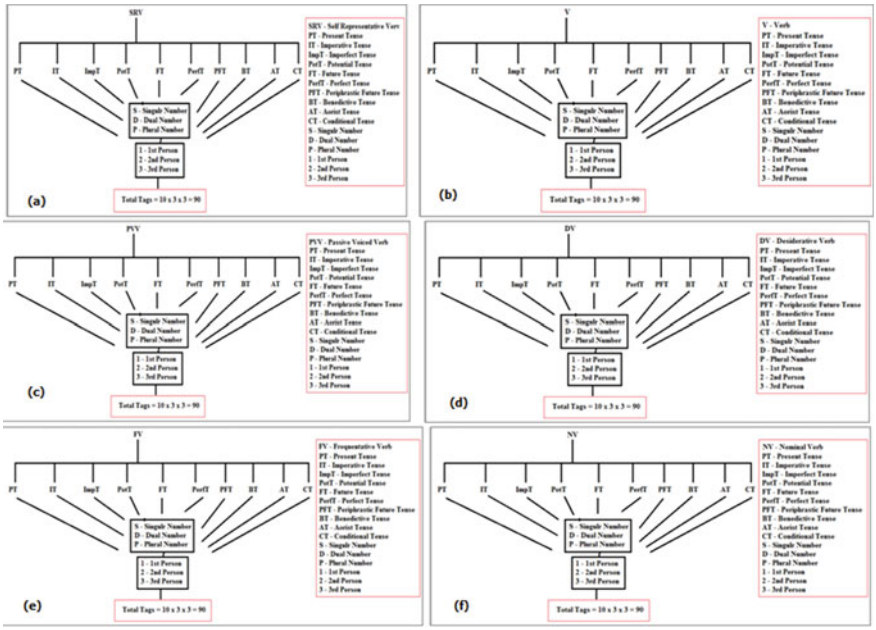


Fig. 2 Verb tags

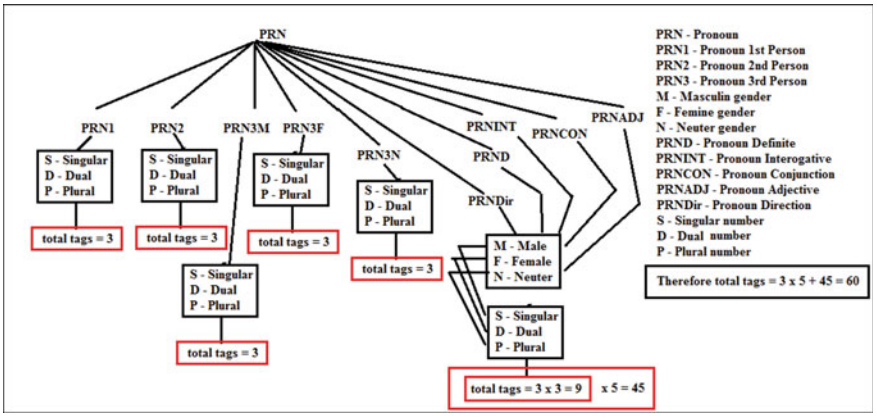


Fig. 3 Pronoun tags

2 Comparative Study of the Previous Work

Table 3 gives a comparison of the previous works with this present work: [2–10].

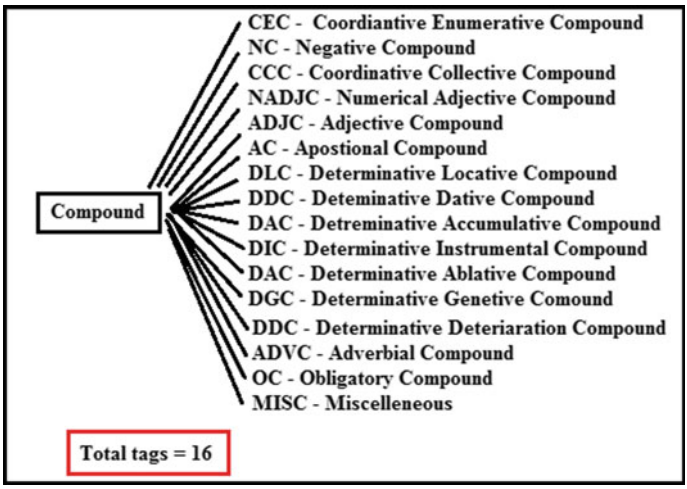


Fig. 4 Compound tags

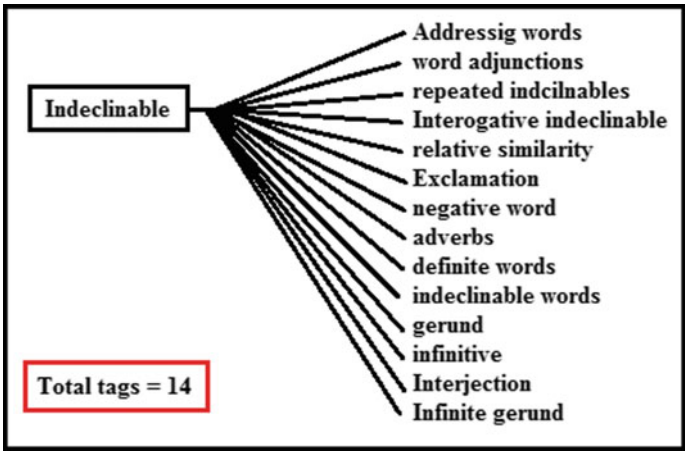


Fig. 5 Indeclinable tags

3 Analysis of Primitive Languages

3.1 Word Order

From Table 4, it is observed that the word order is most flexible in Sanskrit language with 83.33% and then Tamil language with 50%. Though the most of the structural form of used sentences are SOV in Sanskrit and Tamil languages, both these

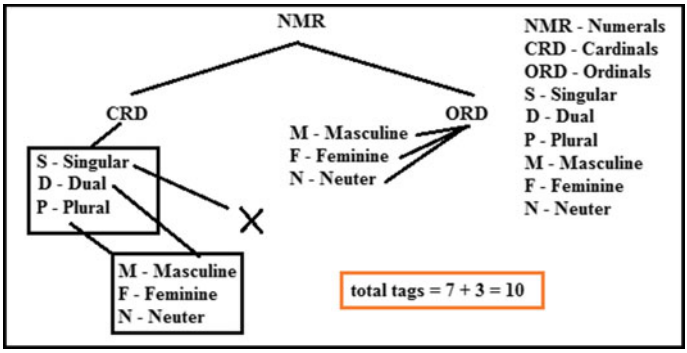


Fig. 6 Numeral tags

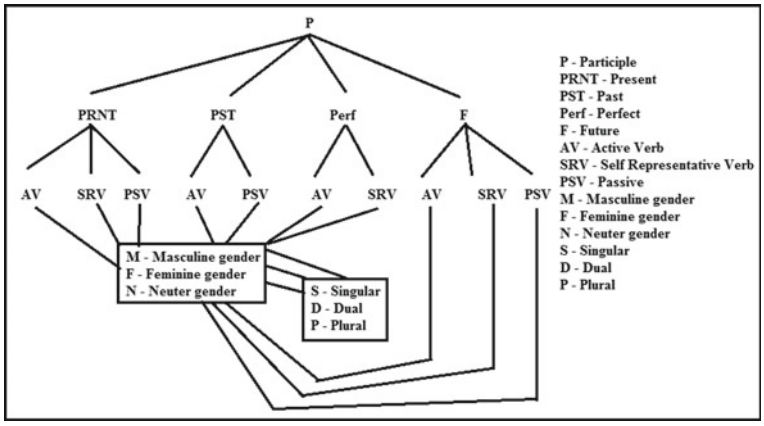


Fig. 7 Participle tags

languages are also compatible in the other word order formats mentioned in the table. But other languages are not compatible with other word order formats.

3.2 The Compound

The compound as defined by the linguistics as a combination of words which are carrying a special meaning. Also it can be defined as a lexeme that consists of more than one stem. The rule for forming a compound is compounding, composition or nominal composition. Table 5 gives the compound for the first five oldest languages.

Table 3 Comparative study of POS in Sanskrit

Development of POS	Institute contribution	Total tags described
JPOS in 2007	JNU	134 (main and sub)
Tirupati POS in 2013	Tirupati University	134 (main)
SAN POS in 2015	Gujrat University	51(main)
IL-POST	Microsoft Research India Lab	28 (main)
Sastra Univ Post for Sanskrit in 2015	Sastra Univ Thanjavar India	11(main)
ILMT	IIIT Hyderabad	26(main)
CPOS IN 2017	JNU	134(main)
Oliver Hellwig in 2016 (new)	Dusseldr of University	136(main)
SANS post in 2019	IIT Kharagpur	130(main)
UPOST in 2022	Royal Global University, Visva Bharati Uinversity and Gauhati University	2023(main and sub)

Table 4 Word order comparison of primitive languages

Languages	No. of POS	SOV	SVO	VSO	VOS	OVS	OSV	WO (%)
Sanskrit	2023	Yes	Yes	Yes	Yes	Yes	No	83.33
Greek	165	Yes	No	No	No	No	No	16.66
Chinese	34	No	Yes	No	No	No	No	16.66
Hebrew	23	No	No	Yes	No	No	No	16.66
Tamil	71	Yes	Yes	No	No	No	Yes	50.00

Table 5 Comparison of compounds in primitive languages

S No.	Name of the languages	No. of compounds
1	Sanskrit	16
2	Greek	9
3	Chinese	5
4	Hebrew	4
5	Tamil	7

3.3 The Basis of Grammar in Primitive Languages

The basis of grammar is based on the factors of a language like (Table 6).

Table 6 Basis of grammar of Sanskrit and other primitive languages

Basis of language	Sanskrit	Greek	Chinese	Hebrew	Tamil
Character set (alphabet)	42	24	Around 6500 mandrains	22	30
No. of vowels	9	5	Not specified	5	12
No. of consonants	33	19	Not specified	17	18
Number	3	2	2	3	2
Order of sentences	5	1	1	1	3
Tense	10	3	Not specified	Not specified	3
Verb mood	4	3	3	2	3
Gender	3	3	Not specified in grammar but preserved in tone	2	2
Part of speech (POS)	2023	165	34	23	71

3.4 The Main and Sub-Tag Classification

In this work, the main tags are classified in its maximum form, and the sub-tags are thoroughly classified, and finally, there are 100 main tags and its associated sub-tags are classified and in results it is found total 2023 tags (Table 7).

Table 7 Main and sub-tag classification of Sanskrit

S No.	POS category	Main POS tags	Sub-POS tags	Total
1	Noun	11	63	693
2	Verb	6	90	540
3	Pronoun	10	63	630
4	Indeclinable	14	–	14
5	Participle	10	9	90
6	Numerical	4	3	10
7	Compound	16	–	16
8	Nipatan	1	–	1
9	Prefix	1	–	1
10	Designation	1	–	1
11	Other language	1	–	1
12	Doubttag	1	–	1
13	Punctuation	25	–	1
Total		100		2023

Table 8 Common and uncommon sub-tags along with their main tags

S No.	POS name	Main tags	Sub-tags				
			Gender	Person	Number	Tense	Case and case-ending
1	Noun	11	X	3	3	X	7
2	Verb	6	X	3	3	10	X
3	Pronoun	10	3	1	3	X	7
4	Indeclinable	14	X	X	X	X	X
5	Participle	10	3	X	3	X	X
6	Numerical	4	3	X	X	X	X
7	Compound	16	X	X	X	X	X
8	Nipatan	1	X	X	X	X	X
9	Prefix	1	X	X	X	X	X
10	Designation	1	X	X	X	X	X
11	Other language	1	X	X	X	X	X
12	Doubttag	1	X	X	X	X	X
13	Punctuation	25	X	X	X	X	X

3.5 The Main Tags with Their Common and Uncommon Sub-Tags Classification

There are 100 main tags are identified in this work and sub-tags are enlisted with the corresponding main tags (Table 8).

4 UPOS Interface Designed for Creating the Database for Sanskrit Language

To create a database for the words of Sanskrit, an interface is designed for preparing the corpus.

The following snapshots are the interface of UPOS (Fig. 8).

5 Observations and Findings

i. According to the evidence, Sanskrit is the oldest language.
ii. The most popular languages fall in the category of closed language, whereas the ancient languages including Sanskrit are open languages.

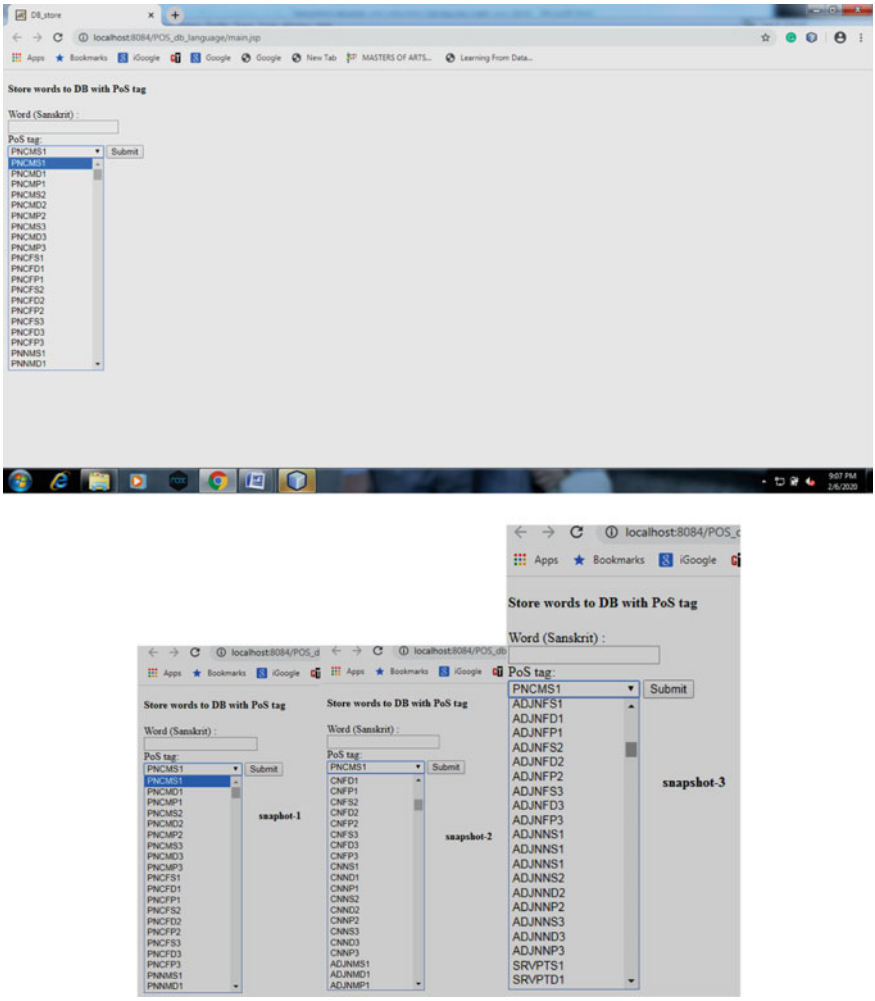


Fig. 8 Snapshots of the interface

iii. The popular languages have boundary-based POS, whereas ancient languages have much more scope of inserting new formation of words in a sentence.
iv. In other previous works, POS tagging: main and its associated sub-features are not discussed except in JNU works.
v. On the other hand, these discussed sub-tags in the previous researches were not clearly analyzed and categorized. As a result, it causes huge confusion in making corpus of the language.
vi. Due to this kind of POS set, word overlapping might occur frequently.
vii. With incomplete POS set, the training data for machine learning will be insufficient, and thus, machine could not perform cent percent.

viii. Word orders in primitive languages are restricted in one or two forms, but in Sanskrit, the word order is more flexible as it supports five forms out of possible six forms.
ix. Sanskrit is considered as the free flow sentence formation language as it supports more than 80% of word order as compared to 50% in Tamil and less than 17% in other primitive languages.
x. The lexeme of Sanskrit language is more because it contains more number of compounds as compared to other oldest or primitive languages.
xi. The communication in human languages is more structured when the time factor or tense is incorporated during their verbal formation of sentences. It is observed in Sanskrit, Greek and Tamil tenses are defined, whereas in Chinese and Hebrew languages, the tense forms are not specified.
xii. Among Sanskrit, Greek and Tamil, Sanskrit presents wide range of tense formation as compared to the Greek and Tamil.
xiii. In Table 7, it is given the more précised form of main and its associated sub-tags.
xiv. Table 7 also identifies more accurately the sub-tags along with their main tags.
xv. Table 7 also helps to prepare the UPOS in more accurate form.
xvi. The most important fact observed here is the sub-tags category case and case-ending which was not discussed in all other previous works.
xvii. To prepare the most efficient corpus of any language depends on the proper definitions and classifications of sub-tags.
xviii. The Sanskrit language can be considered as the most structured language as it contains elaborate classification of main and sub-POS tags.
xix. It is also observed that all the main tags do not have common sub-tags. This phenomenon was not properly shown in the previous works.
xx. Table 7 shows the sub-tags classifications with their main tags. The common and uncommon sub-tags are clearly given in this table.

6 Future Work

Using the database created in this work, a further study will be carried out for making corpus of Sanskrit language which will be a treasure for the researchers of the near future. The corpus will be made available to the world by making them open access to the world.

7 Conclusion

The present work attempted to analyze the primitive languages and developed an interface for making a good database for the Sanskrit language using the new set

of POS called UPOS. The study also pointed out the basic features of the primitive languages of the world. Some interesting findings are observed about the POS set of Sanskrit during the discussion. Hope the present work will give some insight to the primitive languages of the world and also provide an enhance elaborative set of POS set to the linguists.

References

1. https://www.fluentin3months.com/most-spoken-languages/
2. SanskritTagger, a stochastic lexical and POS tagger for Sanskrit. In: Hellwig O
3. A Thesis (2005) The effects of part-of-speech tagsets on tagger performance. In: MacKinlay A, of Melbourne University
4. Hellwig O (2002) Sanskrit und computer. Ph.D. thesis, Freie Universit¨at Berlin
5. Huet G (2007) Shallow syntax analysis in sanskrit guided by semantic nets constraints. In: International workshop on research issues in digital libraries
6. TagMiner: a semisupervised associative POS tagger elective for resource poor languages. In: Rani P, Pudi V, Sharma DM
7. Banko M, Moore RC (2004) Part-of-speech tagging in context. In: Proceeding of COLING
8. Bharati A, Misra Sharma D, Bai L, Sangal R (2006) AnnCorra: annotating corpora guidelines for POS and chunk annotation for indian languages. Tech. Rep. TRLTRC-31, Language Technologies Research Centre, IIIT, Hyderabad
9. Bhatt R, Narasimhan B, Palmer M, Rambow O, Sharma DM, Xia F (2009) A multi-representational and multi-layered treebank for Hindi/Urdu. In: proceeding of the third linguistic annotation workshop, pp 186–189
10. Part-of speech tagger for Sanskrit: a state of art survey. In: Adinarayanan S, Sri Naren N. J Assistant Professor, SASTRA University Thanjavur

A Study on Medical Image Data Augmentation Using Learning Techniques

Vanita D. Jadhav and Lalit V. Patil

Abstract One disadvantage of computer-assisted detection systems is the massive quantity of data needed to train them, which is costly in the medical industry. A big training dataset is critical in deep learning since it enhances training accuracy. Even with a big amount of data, a weak algorithm can be more accurate than a strong algorithm with a little amount of data. When verified on a different unobserved dataset, data augmentation generates new data which is used to train the model and enhances performance. We presented a thorough evaluation of the literature in which data augmentation was employed to train a learning model using lung CT images. Basic and deep learning data augmentation techniques were used to categorize the articles. The term "data augmentation" states a group of approaches for increasing the volume and quality of training datasets. Geometric transformations, kernel filters, color space augmentations, random erasing, mixing pictures, adversarial training, feature space dataset augmentation, meta-learning, and GAN-based networks are among the image augmentation processes explored in this paper. Students will learn how to employ data augmentation to improve model productivity and expand small datasets in order to take advantages of big data.

Keywords Training accuracy · Data augmentation · Big data · Deep learning · Image data · GANs

1 Introduction

In discriminative challenges, Deep Learning representations have achieved unbelievable growth. The development of deep network designs, commanding compute, and access to massive data have all contributed to this. Through the advancement of convolutional neural networks, deep neural nets have been effectively useful to image classification, picture segmentation and object identification. The spatial properties of images are reserved using parameterized, loosely linked kernels in these neural

V. D. Jadhav (✉) · L. V. Patil
SKNCOE, Savitribai Phule Pune University, Vadgaon(Bu), Pune, Maharashtra, India
e-mail: vdjadhav@coe.sveri.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_4

nets. Convolutional layers reduce the 3-D resolution of pictures while increasing the deepness of their feature plots in a sequential manner. This sequence of convolutional changes can produce far lower-dimensional and further useful image illustrations than handmade ones. Convolutional Neural Networks' success has raised interest in using deep learning to solve computer vision issues. The validation error must decline with the training error in order to figure out effective deep learning prototypes. Data augmentation is a very effective way to accomplish this. In this study, we conducted a comprehensive evaluation of the literature in which data augmentation was used to train a deep learning model using lung CT images. Main objective of this study is to prepare live dataset representing CT scan images along with standard dataset and data augmentation methods.

2 Literature Survey

Geometric transformations, feature space augmentation, kernel filters, color space transformations, random erasing, mixing pictures, adversarial training, GAN-based augmentation, meta-learning methods, and neural style transfer are among the augmentations listed in this paper. This segment will describe how each augmentation algorithm works and analyses the method's downsides.

2.1 Geometric Transformations

This section explains how to use geometric transformations to create various augmentation techniques.

a. **Flipping**

A flip is a motion in geometry in which an object is turned over a straight line to form a mirror image. A flip is also called a reflection. Flipping the X axis is more common than flipping the Y axis. It is one of the easier augmentation techniques which has been recognized to work on datasets like ImageNet and CIFAR-10. It is not a label-protective transformation on datasets [1].

b. **Color space**

A tensor of the dimension is commonly used to encode digital image data. Another method which is quite feasible to implement is doing augmentations in the color channels space. Color augmentation can be as simple as separating a single-color channel like R, G, or B. A picture can be rapidly transformed into another picture in one color channel by dividing a matrix and addition of two zero matrices from the additional color stations. Using basic matrix operations, RGB values can be simply changed to enhance or lower the image's brightness.

Adjustments are made by altering the intensity levels in the histograms, similar to those found in photo editing software [2].

c. **Cropping**

Cropping photographs with varied height and width proportions are used as a valued processing step for image data. Furthermore, Random cropping can be used to create a similar effect to translations [1]. Random cropping varies from translations as it decreases the size of the input, whereas translations keep the image's 3-D dimensions.

d. **Rotation**

The image can be rotated right or left on an axis between 1° and 359° in rotation augmentations. The rotation degree parameter has a great influence on the safety of rotation augmentations [2].

e. **Translation**

Continuously shifting images right, left, down, and up can be a highly beneficial alteration for avoiding data positional bias. For instance, if all images in a dataset are perfectly positioned, the model must be validated on such pictures. The leftover space can be occupied with a constant value for example 0 s or 255 s, when the original picture is translated in a direction. This filling keeps image's 3-D dimensions after it is been augmented.

f. **Noise injection**

It is the process of inserting a matrix of arbitrary values, generally derived from Gaussian distribution. Addition of noise to photos can aid CNNs in recognizing further distinct features [3]. An excellent approach to cope with training data positional biases is Geometric transformation. There are a variety of biases that might cause training data distributions to deviate from testing data distributions. Geometric transformations are also advantageous because they are simple to apply. Increased memory, additional training time, and transformation computation costs are some of the drawbacks of geometric transformations [4].

2.2 Color Space Transformations

The picture data is divided in 3 matrices, which has a different size. These matrices signify the pixel values for each RGB color. Igniting biases are one of the most common problems that image recognition face. As a result, determining the efficiency of color space modifications (photometric transformations) is quite straightforward. Twisting over the photos and reducing or growing the pixel values by a fixed value is an easy solution for very bright or shady images. Another simple color space transformation is merging out individual RGB color matrices. Limiting pixel values to a defined least or extreme value is another adjustment. The inherent

representation of color in digital photographs allows for a wide range of augmentation approaches. This transformation can also be used in image-editing software [5]. Converting RGB matrices into a single grayscale image simplifies the representation of picture datasets. There are various ways to express digital color, such as HSV, besides RGB versus grayscale photos [6].

Color space conversions have several drawbacks, including bigger memory, cost of transformation, and time required for training. Because color modifications might potentially remove important color information, they are not necessarily a label-preserving alternative [7].

2.3 Kernel Filters

These are a type of image processing technique that can be used to improve and shape images. These filters use Gaussian blur filter to move a nxn matrix across an image, resulting in sharpy image along the edges. When pictures are blurred on the fly for data augmentation, they may be more resistant to gesture blur in testing. Furthermore, when images are refined for data augmentation, additional details about things of interest may be recorded. Sharpening and blurring are two examples of how kernel filters can be used on pictures. Kernel filters function better as a network layer rather than as a data augmentation dataset addition [8].

2.4 Mixing Images

A method of data augmentation that involves combining images by be around their pixel values is paradoxical. To a human viewer, the visuals created by this method will not appear to be a useful alteration. Another finding of the study is that when photos from the complete training set were mixed instead of instances from the same class, better results were obtained [9, 10].

This strategy has the obvious disadvantage of making little sense from a human standpoint. It is tough to comprehend or express the performance improvement that comes from combining images. One possibility is that as the dataset size grows, low-level properties like lines and edges become more robustly represented. The act of this strategy in comparison to pertaining methods and transfer learning is an intriguing field for further research. Other strategies for learning low-level properties in CNNs include transfer learning and pertaining [11, 12].

2.5 Random Erasing

It is a data augmentation technique invented by Zhong et al. Random erasing is related to dropout regularization, and it is based on dropout regularization mechanics. This method was developed to overcome occlusion-related image identification challenges. When some sections of an object are obscured, this is referred to as occlusion. Random erasure prevents this through encouraging the model to acquire additional graphic features of image, avoiding it from becoming fixated on a single graphic component [13]. Random erasing is a potential approach for guaranteeing that a network considers entire image instead of just a section of image, aside from the occlusion visual problem. Random erasing selects nxm pixels in a picture at random and masks them with 0 s, 255 s, random values or mean pixel values [14]. Additional augmentation methods, like color filters or horizontal flip, can be built on top of this augmentation method. Random erasing has the drawback of not necessarily being a label-preserving change. [15].

2.6 Feature Space Augmentation

All the above augmentation approaches are used on images in the input space. It is particularly impressive that neural networks can transform high-dimensional inputs into lower-dimensional pictures. In fattened layers, neural networks can map pictures to binary classes. Neural network's intermediate representations can be isolated from the network as a whole by altering the network's sequential processing. It is possible to separate and isolate the lower-dimensional pictures of visual input in fully-connected layers. The lower-dimensional pictures contained in the high-level layers of a CNN are stated to as the feature space. SMOTE is a well-liked augmentation for fixing concerns of class imbalance. By combining the K-nearest neighbors, this method is utilized to build new instances in the feature space. Feature space augmentation can also be accomplished by separating vector images from a CNN [16]. It is accomplished by slicing the network's output layer, resulting in a low-dimensional vector as the output instead of a class label. In the future, the effectiveness of this technique will be examined further [17]. A difficulty of feature space augmentation is that it is very difficult to interpret the vector data.

2.7 Adversarial Training

This is a method of using 2 or more than two networks with loss functions that have contrasting purposes set in them. Noise search or augmentation search which is application of adversarial training is still a new notion that has not been extensively explored. Though it has been demonstrated that employing adversarial search to insert

noise increases performance on adversarial instances, it is unknown whether this is also effective in reducing overfitting. The link between adversarial attack resistance and actual performance on test datasets will be the focus of future research [18].

2.8 GAN-Based Data Augmentation

Generative modeling is additional fascinating data augmentation approach. The exercise of constructing artificial instances from a dataset with parallel features to the unique set is known as generative modeling. The GAN-based data augmentation context can be expanded to advance the excellence of auto-encoder models. The outstanding performance of GANs has sparked a renewed interest in how they may be used for data augmentation. Neural networks are having the capability to provide more training data, resulting in more accurate classification models. GANs have disadvantage of requiring a big quantity of data to train [19].

2.9 Neural Style Transfer

It is one of the most fascinating demonstrations of deep learning skills. Fundamental impression is to use CNN-generated image representations. Although it is most known for its aesthetic uses, Neural Style Transfer can also be utilized for data augmentation. The approach manipulates the subsequent images in a CNN in such a way that the style of one image can be transferred to another while the original content is preserved [20].

A quick taxonomy of the data augmentations is shown below in Fig. 1.

3 Conclusion

This study classifies data augmentation hypotheses for the occurrence of overfitting in DL models due to a shortage of data. To avoid overfitting, DL models trust on large amounts of data. The benefits of large data in the restricted data realm can be achieved by precisely expanding datasets by means of the approaches mentioned in this survey. Data augmentation is a powerful tool for improving dataset quality. Deep neural networks' layered architecture opens up a lot of possibilities for data augmentation. The input layer is where the majority of the augmentations surveyed function. Some, however, are generated from hidden layer representations. The label space and the space of intermediate representations are two not yet explored areas of data augmentation with promising outcomes. Although many of these approaches and principles can be applied to other data domains, this study concentrates on applications for medical picture data. Data augmentation has a bright future ahead of

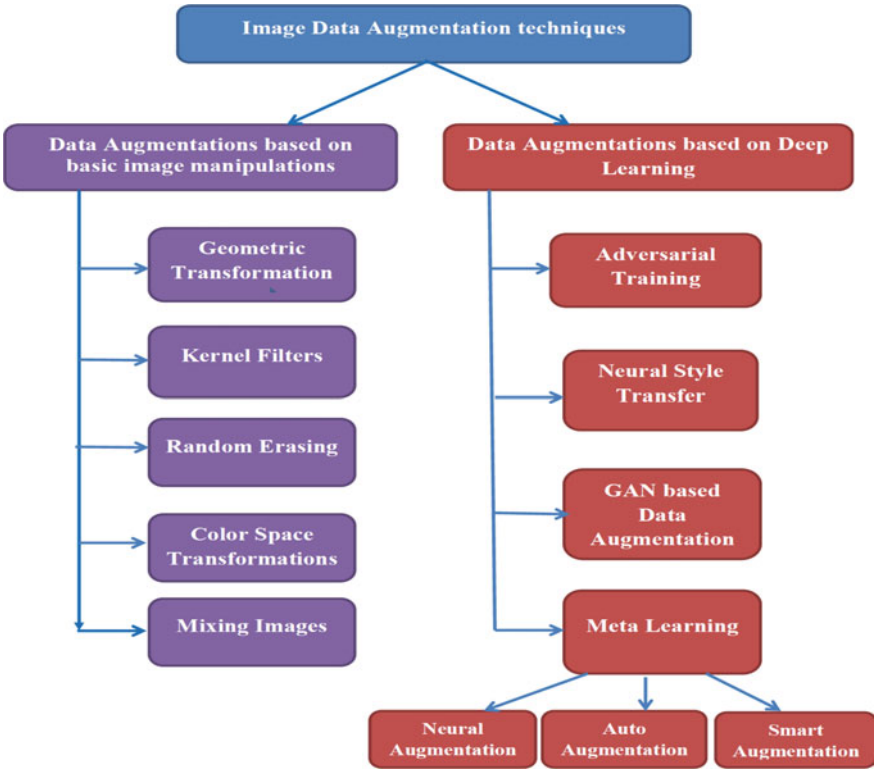


Fig. 1 Taxonomy of the data augmentations

it. The potential for using search algorithms that combine data warping and over-sampling methods is immense. Deep neural networks' layered architecture opens up a lot of possibilities for data augmentation. Main objective of this study is to prepare live dataset representing CT scan images along with standard dataset and data augmentation methods.

References

1. Olaf R, Philipp F, Thomas B (2015) U-Net: convolutional networks for biomedical image segmentation. In: MICCAI. Springer, p 234–41
2. Ekin DC, Barret Z, Dandelion M, Vijay V, Quoc VL (2018) Auto augment: learning augmentation policies from data. ArXiv preprint
3. Francisco JM-B, Fiammetta S, Jose MJ, Daniel U, Leonardo F (2018) Forward noise adjustment scheme for data augmentation. ArXiv preprints
4. Dua D, Karra TE (2017) UCI machine learning repository. University of California, School of Information and Computer Science, Irvine, CA. http://archive.ics.uci.edu/ml

5. Ken C, Karen S, Andrea V, Andrew Z (2014) Return of the devil in the details: delving deep into convolutional nets. In: Proceedings of BMVC
6. Mark E, Luc VG, Christopher KIW, John W, Andrew Z (2008) The pascal visual object classes (VOC) challenge. http://www.pascal-network.org/challenges/VOC/voc2008/workshop/
7. Aranzazu J, Miguel P, Mikel G, Carlos L-M, Daniel P (2010) A comparison study of different color spaces in clustering base image segmentation. IPMU
8. Guoliang K, Xuanyi D, Liang Z, Yi Y (2017) Patch Shuffle regularization. arXiv preprint
9. Hiroshi I (2018) Data augmentation by pairing samples for images classification. ArXiv eprints
10. Cecilia S, Michael JD (2018) Improved mixed-example data augmentation. ArXiv preprint
11. Daojun L, Feng Y, Tian Z, Peter Y (2018) Understanding mixup training methods. In: IEEE access
12. Ryo T, Takashi M (2018) Data augmentation using random image cropping and patches for deep CNNs. arXiv preprints
13. Zhun Z, Liang Z, Guoliang K, Shaozi L, Yi Y (2017) Random erasing data augmentation. ArXiv e-prints
14. Terrance V, Graham WT (2017) Improved regularization of convolutional neural networks with cutout. arXiv preprint
15. Agnieszka M, Michal G (2018) Data augmentation for improving deep learning in image classification problem. In: IEEE 2018 international interdisciplinary Ph.D. Workshop
16. Tomohiko K, Michiaki I (2018) Icing on the cake: an easy and quick post-learning method you can try after deep learning. arXiv preprints
17. Terrance V, Graham WT (2017) Dataset augmentation in feature space. In: Proceedings of the international conference on machine learning (ICML), workshop track
18. Seyed-Mohsen MD, Alhussein F, Pascal F (2016) DeepFool: a simple and accurate method to fool deep neural networks. arXiv preprint
19. Jiawei S, Danilo VV, Sakurai K (2018) One pixel attack for fooling deep neural networoks. arXiv preprints
20. Leon AG, Alexander SE, Matthias B (2015) A neural algorithm of artistic style. ArXiv

Data Classification Using Mesh Generation and Hough Transform

Abdoulaye Sere, Harrisson Thiziers Achi, Jacques Rodrigue Guiguemde, Tiemoman Kone, and Kisito Kabore

Abstract Linear regression is a well-known method used in statistic to determine correlation between two or more variables, in machine learning for data classification and for prediction. This paper deals with data classification using a composition of two techniques of classification such as mesh generation, followed by the Hough transform method, in order to define a linear regression on the dataset. The purpose is to use analytical straight lines as a regression technique. The proposed method is a technique of data reduction through a mesh generation, which creates a virtual grid where each continuous point is localized in a cell corresponding to a pixel in the virtual grid. The standard Hough transform method establishes a relation between an image space and a parameter space through the definition of a sine function. The values of variables are represented by the coordinates of continuous points in the image space. The standard Hough transform is applied to each cell. The coordinates parameters with the high number of votes superior to a threshold that appears in the parameter space or the accumulator correspond to the parameters of analytical straight lines. Straight line recognition achieves linear regression and establishes correlation between the initial variables. Our analysis presents difference between classical linear regression and the proposed alternative which gives more visibility between data relation and accept a level of error in the dataset.

A. Sere (✉) · J. R. Guiguemde · K. Kabore
Equipe de Recherche Signal, Image et Communications (ERSIC) Laboratoire d'Algèbre de Mathématiques Discrétes et Informatique (LAMDI), Université Nazi BONI, Houet, Burkina Faso
e-mail: abdoulaye.sere@recifaso.org
URL: https://recifaso.org

A. Sere · J. R. Guiguemde
Réseau des Enseignants Chercheurs et Chercheurs en Informatique du Faso (RECIF), Ouagadougou, Burkina Faso

H. T. Achi · T. Kone
Unité de Recherche et d'Expertise Numérique (UREN), Université Virtuelle de Côte d'Ivoire, Yamoussoukro, Côte d'Ivoire
e-mail: thiziers.achi@uvci.edu.ci

T. Kone
e-mail: tiemoman.kone@uvci.edu.ci

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_5

Keywords Hough transform · Classification · Machine learning

1 Introduction

In discrete geometry, several definitions of digital straight lines and digital circles have been defined by scientists to draw lines and circles in an image. For instance, famous Bresenham's line and Bresenham's circle in [1, 2] have been classical algorithms, very known in this ways. Then, analytical straight lines have been also proposed by Reveillès in [3], Andres in [4] and Sere in [5]. Many works have carried out different techniques of data classification used in machine learning, as decision tree, fuzzy classification, neural network, linear regression and support vector machine (SVM). Linear regression and support vector machine (SVM) are based on straight line definition. The questions are:

- how to extend linear regression, in considering analytical straight line definitions proposed by Reveilles in [3], Andres in [4] and Sere in [5]?
- in this way, how to proceed by the standard Hough transform method for data classification?

Since 1962, Paul Hough in [6] introduced the Hough transform method to detect straight lines. Several extensions and applications have been studied by scientists for recognition of analytical straight lines as the generalized preimage presented by Dexet's thesis in [7] and the dual of a square defined in Sere's thesis [8] with its applications in [9].

Moreover, the Hough transform method allows recognition of others geometric shapes as parameterized curves, circles, ellipse and arbitrary shapes that have been proposed by Duda and others in [10]. Many applications of Hough transform have been realized in mouth recognition in [11] and action recognition in [12].

Many extensions of the Hough transform method have been realized by Dexet and others in [13] to compute the dual of a pixel and a voxel, by Sere and others in [14] to compute the dual of a square and others geometric shapes as triangles, rectangles, octagons and hexagons, following the standard Hough transform definition in [5, 9, 15].

This paper introduces a new technique of data classification through linear regression based on the standard Hough transform method. Our purpose is precisely to introduce analytical straight line definitions as a linear regression technique and to establish its detection by the Hough transform method.

Therefore, this paper is organized as follows: Sect. 2 recalls the definition of analytical straight lines, standard Hough transform and different concepts related to linear regression, useful in the best understanding of the next sections. Section 3 exposes the method based on function composition, how tiling the space with a rectangular grid using the parameters α, β to measure the rectangle sides, followed by the application of Hough transform. Experimental results illustrated in Sect. 4 compare the implementation of the proposed method and classical linear regression.

2 Preliminaries

This section is focusing on regression definition, the Hough transform method and analytical straight line that will be used forward for linear regression construction and detection. This section contributes to the best understanding of the following sections, namely method description and simulation.

2.1 Linear Regression

Regression is a well-known statistical method used in finance, investing, health, economy and machine learning for prediction and others disciplines, to determine the relationship between one dependent variable, denoted Y and a series of others variables, denoted X_i.

The two basic types of regression are simple linear regression and multiple linear regression. But, there are also nonlinear regression methods, used for analysis on more complicated data. Others types of regression as polynomial function and splines give possibilities to determine relation in the dataset.

Simple linear regression uses one independent variable to explain or to predict the outcome of the dependent variable Y, while multiple linear regression uses two or more independent variables to predict the outcome.

We recall definitions related to these basic regressions:

Definition 1 (*Simple Linear Regression*) Let X, Y be two variables. Simple linear regression between X and Y is defined by the general relation: $Y = a + bX + u$ where *a, b, u* are constants and $X = (x_i)$, $Y = (y_i)$ with $i \in \{1, 2, \ldots, n-1, n\}$

That means, each pair (x_i, y_i) respectively from the pair (X, Y), verify the relation:

$$y_i = a + bx_i + u. \tag{1}$$

Definition 2 (*Multiple Linear Regression*) Let X_i, Y be variables where $i \in \{1, 2, ..., n-1, n\}$. Multiple linear regression is defined by the general relation $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n + u$ where a, b_k, u are constants with $k \in \{1, 2, \cdots, n-1, n\}$ and $Y=(y_j)$, $X_i=(x_{i,j})$

That leads to the relation:

$$y_j = a + b_1x_{1,j} + b_2x_{2,j} + b_3x_{3,j} + \cdots + b_nx_{n,j} + u \tag{2}$$

Suppose a dataset with the variables X and Y representing, respectively, temperature data and maximal concentration of ozone. For instance, Fig. 5, in appendix

illustrates a linear regression for the dataset represented by black points (x_i, y_i) where X=(x_i) and Y=(y_i).

The relations 1 and 2 are easily represented by artificial neural networks, in considering x_i as inputs with ponderation parameter a, b_k and the output function y_i adding u. There are similarity between regression and the definition of Hough transform extension introduced in Dexet's thesis in [7], where the dual of a continuous point, the dual of a pixel and the dual of a voxel have been studied. For more details, we have:

$$y_i = a + bx_i + u \Longleftrightarrow u + a = y_i - bx_i \tag{3}$$

Equations 3 correspond to the dual of a continuous point (x_i,y_i) in 2D in the accumulator, according to Martine Dexet in [7]. Moreover, Eq. 2 becomes:

$$u + a = y_j - nb_1x_{1,j} - b_2x_{2,j} - b_3x_{3,j} - \cdots - b_nx_{n,j} \tag{4}$$

That also means the dual of a continuous point $(x_{1,j}, x_{2,j}, \ldots, x_{n,j},y_j)$ in n+1 dimensional space in the parameter space (accumulator), according to Martine Dexet in [7].

Therefore, regression is equivalent to classical Hough transform method. Thus, Spratling in [16] has even explained the application of Hough transform through the neural networks. In [17], Bamogo evaluates the impact of Hough transform and the neural network on recognition of straight lines: the preprocessing of images by Hough transform for detection of straight lines before the neural network application improves detection rates.

The question is why not to consider Hough transform extension as a regression technique? Here, we are focusing on the application of the standard Hough transform introduced in Sere's thesis in [8], in considering analytical straight line recognition, as a regression technique.

2.2 Analytical Straight Line

Analytical straight line is defined by Reveillès in [3], extended by Andres in [4] to define supercover model. Sere and others in [5] also introduced new analytical straight lines that give another's alternatives to draw analytical straight lines in a hexagonal grid or in a octagonal grid.

Definition 3 (*Analytical straight line*) analytical digital straight line [3, 4] with parameters (a, b, μ) and thickness w is defined by the set of integer points (x, y) verifying: $\mu \leq ax + by < \mu + w$, $(a, b, \mu, w) \in Z^4$, $pgcd(a, b) = 1$.

Thus, analytical digital straight line is then:

- thin if $w < \max(|a|, |b|)$
- naif if $w = \max(|a|, |b|)$

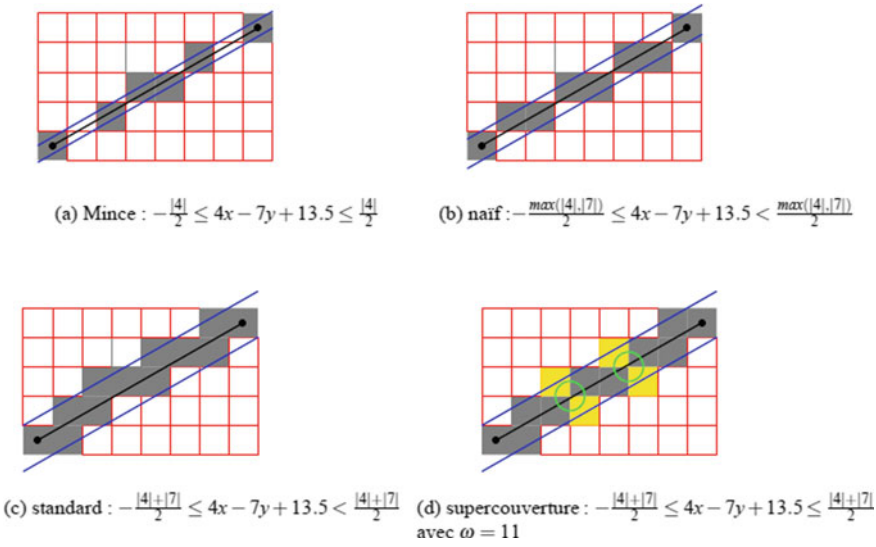


Fig. 1 Examples of analytical straight lines, successively thin, naif, standard, supercover from Sere's thesis in [8] that could be used as regression techniques in our analysis

- standard if $w = (|a| + |b|)$
- thick if $w > (|a| + |b|)$.

For instance, Fig. 1 from Sere's thesis in [8] illustrates thin, naif, standard and supercover analytical straight lines.

Our idea is that thin, naif, standard and supercover analytical straight lines can be used as linear regression, through the application of extended Hough transform proposed by Dexet in [7] and Sere in [14] to recognize analytical straight line.

2.3 Standard Hough Transform

Standard Hough transform (SHT) in [8] creates a relation between two spaces, an image space and a parameter space. It establishes a sine function in the parameter space as follows:

Definition 4 (*Standard Hough Transform (SHT of a point)* [8]) The dual *S(x,y)* of the point (*x, y*) in a two-dimensional space where $(x, y) \in R^2$, is the standard Hough transform defined by the set of points:

$$\{(\theta, r) \in [0, \pi] \times [-\sqrt{l^2 + h^2}, \sqrt{l^2 + h^2}] / \; r = x \cos\theta + y \sin\theta\} \qquad (5)$$

Standard Hough transform has been extended by Sere and others in [8] to define its application to geometry shapes such as squares, triangles, octagons and hexagons.

For example, the application of standard Hough transform to a square is the union of all the points localized between the dual of its internal diagonals, accordingly to the works, presented by Sere and others in [14].

3 Method Description

This section concerns with the description of our method that combines data classification and Hough transform to establish linear regression: it uses precisely a tiling space and the standard Hough transform (denoted SHT) to achieve data classification. Tiling space gives possibility to put together data in the same cluster corresponding to generated pixels. Our idea is to analyze a cloud of continuous points through a grid where a set of discrete points are used to establish linear regression. The coordinates of continuous points are real while those of discrete points are only integer.

Let $X = \{X_1, X_1, \ldots, X_{n-1}, X_n\}$ be a set of n points such as $X_i \in W$, where W is the universal set of points. The set X is concerned with Linear Regression. X_i is a vector and takes several instances in a n dimensional space, as defined by the tuple $X_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,n-1}, a_{i,n})$. The tuple $(a_{i,1}, a_{i,2}, \ldots, a_{i,n-1}, a_{i,n})$ is transformed into a discrete tuple to get integer coordinates to have the nearest cells as follows: $b_{i,j} = E(a_{i,j})$ where $E(x)$ is the integer approximative function of $x : E(x)$ returns the nearest approximative integer value of x.

In this manner, each $X_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,n-1}, a_{i,n})$ is associated to $Y_j = (b_{j,1}, b_{j,2}, \ldots, b_{j,n-1}, b_{j,n})$ as the corresponding approximative integer values.

Finally we have:

$$[E(X_i) = Y_i] \Longleftrightarrow [b_{i,j} = E(a_{i,j})] \tag{6}$$

A fixed cluster $Y = (b_1, b_2, \ldots, b_{n-1}, b_n)$ is defined by a particular set of X_i where i={k,…, l} and {k,…, l} ⊂ {1,2,…, n-1, n } such as:

$$[E(X_i) = Y] \Longleftrightarrow [b_j = E(a_{i,j})] \tag{7}$$

That means, there exist generally the parameters α_i such as all these conditions are true:

$$|b_1 - a_{i,1}| \leq \alpha_1 \tag{8}$$

$$|b_2 - a_{i,2}| \leq \alpha_2 \tag{9}$$

$$|b_{n-1} - a_{i,n-1}| \leq \alpha_{n-1} \tag{10}$$

$$|b_n - a_{i,n}| \leq \alpha_n \tag{11}$$

The proposed method consists of building a composition of functions: that means using $E(x)$ followed by the standard Hough transform (SHT), denoted by the formula SHT $oE(X_i)$.

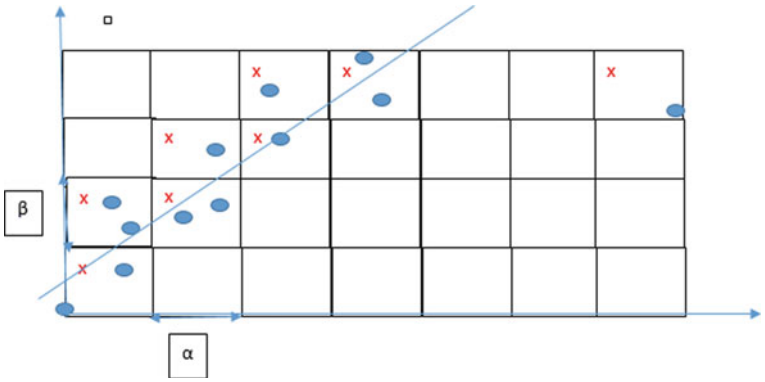


Fig. 2 Cluster and linear regression

Therefore, we introduce a grid in the initial image space to classify continuous points into different cells as illustrated in Fig. 2. Each continuous point in the image space belongs to a cell, representing all the continuous points inside.

The size of different cells or cluster influes on data classification to determine linear regression. In a two-dimensional space, a rectangular grid leads to clusters Y_i defined by $Y_i = (b_{i,1}, b_{i,2})$ with $X_i = (a_{i,1}, a_{i,2})$. The parameters α_1 and α_2 become respectively α, β corresponding to the size of different sides for the cluster (or the rectangle). α and β depends on relevance to put X_i together to respect distance restriction $|b_{i,1} - a_{i,1}| \leq \alpha$ and $|b_{i,2} - a_{i,2}| \leq \beta$.

How to proceed to mesh generation in a two-dimensional space? Consider R be a rectangular area. Let a, b, c, d be respectively $a = \min(a_{i,1})$, $b = \max(a_{i,1})$, $c = \min(a_{i,2})$ and $d = \max(a_{i,2})$. The points of coordinates (a, c), (a, d), (b, c), (b, d) describes in a two-dimensional space, the vertices of the rectangular area R that contains all the points $X_i = (a_{i,1}, a_{i,2})$ where i $\in \{1, 2, \ldots, n-1, n\}$. The parameters α, β are defined by:

$$\begin{cases} \alpha = \frac{|b-a|}{k} \\ \beta = \frac{|d-c|}{l} \end{cases} \tag{12}$$

Where k, l are integer. Therefore, when k and l are fixed, we will obtain easily the values of α and β.

This method tolerates errors on each X_i in the limited surface $(\alpha. \beta)$, represented by the rectangular cluster.

The dual of clusters that contain a number of continuous points inside superior to the threshold γ could be computed in the accumulator to reduce the number of dual, as an optimization.

Figure 2 illustrates an example of a cloud of continuous points in a two-dimensional space. Extremity points are used to build the rectangular area with different clusters that contain the values X_i.

Linear regression leads to apply Hough transform to each cluster: it gives a set of straight lines, having the high vote number in the accumulator. These straight lines pass through different clusters.

There are two families of straight lines, vertical straight line and horizontal straight line, following the definition of quasi affine application:

- the family of vertical straight lines is given by the relation $x = t.\ \alpha$ where $t \in \{0, 1, 2, \ldots, k\}$
- while the family of horizontal straight lines is defined by the egality $y = v.\ \beta$ where $v \in \{0, 1, 2, \ldots, l\}$

Thus, each pair (x_i, y_i) is transformed to the pair (t, v) of integer in a new repere, as follows:

$$\begin{cases} t.\alpha \leq x_i \leq (t+1).\alpha \\ v.\beta \leq y_i \leq (v+1).\beta \end{cases} \tag{13}$$

The constraints in the relation 13 lead to a rectangle, having its side length defined by α, β and its center represented by integer coordinates (t, v), associated to all the pairs (x_i, y_i).

The dual of a rectangle (or cluster) will be computed in the accumulator, if the number of continuous points (x_i, y_i) inside, is superior to the threshold γ. This last restriction also represents a fuzzy set model to select the list of valid clusters, being considered for the dual in the accumulator. For instance, Algorithm 2 in appendix illustrates selection and computing the dual of valid clusters.

In the accumulator, it appears a set of cells having the high vote number. Consider (θ_1, r_1) an example of cell. We have the relation r_1= x cos (θ_1) + y sin (θ_1).

Suppose $\mu \leq ax + by < \mu + w$ with a=h cos (θ_1), b= h sin (θ_1) and μ = h r_1 where h is a real constant, is an analytical straight line, found in the new repere, accordingly to the highest vote in the accumulator. We have:

$$\mu\alpha\beta \leq a\alpha\beta x + b\alpha\beta y < (\mu + w)\alpha\beta \tag{14}$$

That means with X= αx and Y= βy in the original repere:

$$\mu\alpha\beta \leq a\beta X + b\alpha Y < (\mu\alpha\beta + w\alpha\beta) \tag{15}$$

Suppose X is fixed. Linear regression based on the standard Hough transform determines an interval for the variable Y, as follows.

If b>0, then:

$$\frac{\mu\alpha\beta - a\beta X}{b\alpha} \leq Y < \frac{(\mu\alpha\beta + w\alpha\beta) - a\beta X}{b\alpha} \tag{16}$$

else ($b < 0$):

$$\frac{(\mu\alpha\beta + w\alpha\beta) - a\beta X}{b\alpha} \leq Y < \frac{\mu\alpha\beta - a\beta X}{b\alpha} \tag{17}$$

Moreover, in mesh generation, others geometry shapes, as triangles, hexagons and octagons lead to establish new models to represent clusters. In this manner, standard Hough transform of these new cells is evaluated in the accumulator in order to get parameters for analytical linear regression.

4 Simulation and Discussions

This section concerns with simulation of the proposed method through a program in Python. Consider a given dataset of one hundred pairs related to making relation between declared turnovers and corresponding tax payment for anonymous enterprises in Burkina Faso, collected by General Direction of Tax, as mentioned in the Table 1 with a sample of ten extracted pairs (turnovers, tax payment).

Fig. 3, in appendix represents one hundred of pairs related to turnovers and tax payment in a repere. It also illustrates three groups of pairs according to continuity of points:

- one separated point (more on the right) that could be wrong data;
- two separated points on the top;
- a group of continuous points on the bottom.

The separated point on the right has been eliminated in the data and the scale of the repere has been decreased, as illustrated in Fig. 4, in appendix: it appears more informations in the initial group which is divided into several sub-groups. Therefore, multi-scale analysis with variation on scale highlights details on proximity and relations between data. That shows effectively importance of playing on α and β

Table 1 A sample of ten extracted pairs (turnovers, tax payment) from the dataset

Declared turnover (x_i)	Corresponding tax payment (y_i)
828,119,285	25,655,100
600,421,484	23,612,142
197,996,735	6,946,016
141,416,533	611,073
70,559,937	1,197,763
67,360,002	2,344,750
54,147,500	2,271,249
53,780,081	6,395,359
50,599,898	7,068,074
49,478,528	12,934,269

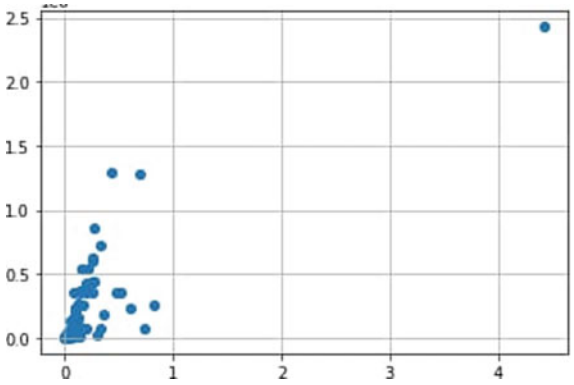


Fig. 3 Representation of one hundred pairs (turnover, tax payment in initial data

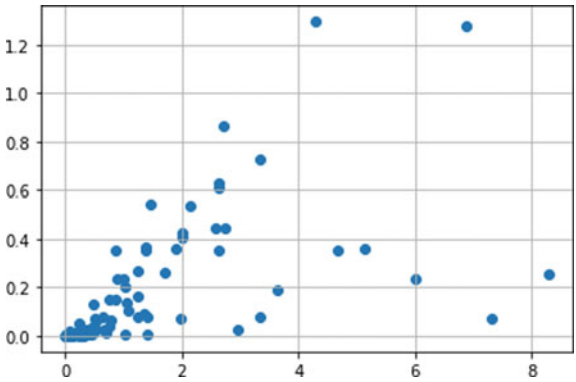


Fig. 4 Representation of one hundred pairs (turnover, tax payment in a decreased scale

respectively for each axis to view proximity between different points in the image space.

Formerly, α and β lead to establishment of 4-neighborhood or 8-neighborhood in the dataset. A point, said to be far from another's points depends on the scale defined by the value of α and β.

For all the data (one hundred pairs), particularly for data in Table 1, the coordinates of the rectangle vertices (a, c), (b, c), $(a, d)(b, d)$, as previously specified in Sect. 3, are defined by: $a = 65181778$, $b = 828119285$, $c = 7549914$ and $d = 129468861$.

Our analysis presents two ways:

- the first alternative in Sect. 4.1 computes the average of $|a_{j,1} - a_{i,1}|$ and $|a_{j,2} - a_{i,2}|$, associated respectively to α and β with i $\neq$ j, only with the initial data (ten rows) in Table 1.
- the second one in Sect. 4.2 is based on three cases of arbitrary pairs (l, k) as parameters.

Table 2 Data analysis with the average

Declared turnovers (x_i)	Tax payment (y_i)	cluster x (E($\frac{x_i}{\alpha}$))	cluster y(E($\frac{y_i}{\beta}$))
828,119,285	25,655,100	3	3
600421484	23,612,142	2	2
197996735	6,946,016	1	1
141416533	611,073	1	0
70559937	1,197,763	0	0
67360002	2,344,750	0	0
54147500	2,271,249	0	0
53780081	6,395,359	0	1
50599898	7,068,074	0	1
49478528	12,934,269	0	1

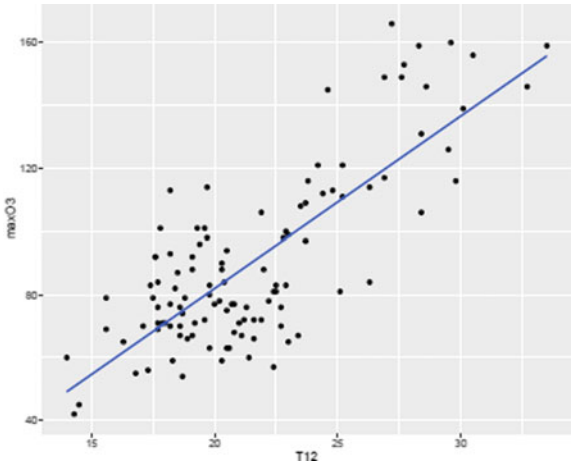


Fig. 5 Linear regression with dataset (X, Y), representation of temperature at twelve oclock (T12) in x-axis and maximal concentration of ozone (maxO3) in y-axis

4.1 Computing Average of $|a_{j,1} - a_{i,1}|$ and $|a_{j,2} - a_{i,2}|$ with $i \neq j$

The values of α and β are respectively defined by the relation: $\alpha = \frac{\Sigma|a_{j,1} - a_{i,1}|}{N}$ and $\beta = \frac{\Sigma|a_{j,2} - a_{i,2}|}{N}$ where N is the number of pairs (i, j) such as $i \neq j$.

As $N = C_p^2$ with $i, j \in \{1, 2, \ldots, p\}$, we have: $\alpha = \frac{\Sigma|a_{j,1} - a_{i,1}|}{C_p^2}$ and $\beta = \frac{\Sigma|a_{j,2} - a_{i,2}|}{C_p^2}$

Algorithm 1 in appendix, computes the values of α and β. For instance, only with the initial data (ten rows) in Table 1, we have computed α = 263169071,5 and β = 10007391,67. These data is also represented in Fig. 6 in appendix, also corresponding

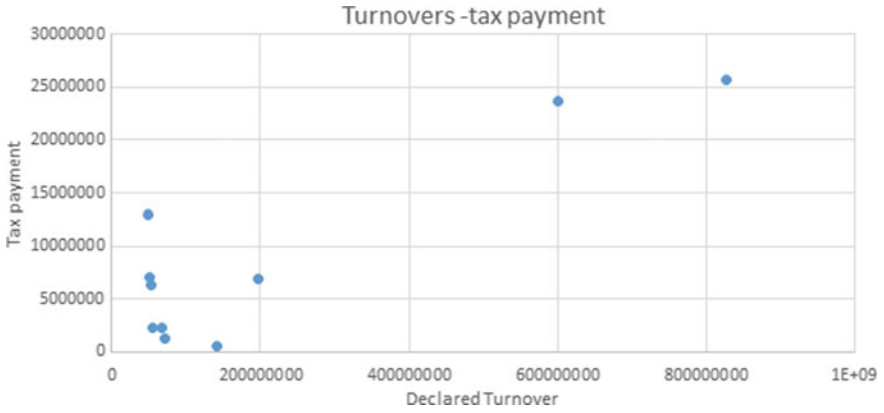


Fig. 6 Representation of ten continuous points, where data are from the first two columns of Table 2, in appendix

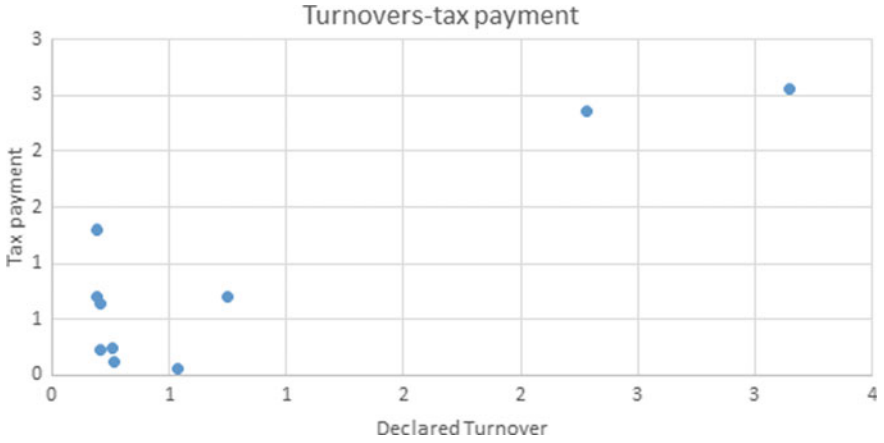


Fig. 7 Representation of ten continuous points, where data are from the first two columns of Table 2, in appendix in the new repere, in application of α and β to each axis

to the first two columns of Table 2 in appendix. These continuous points are also represented in a new repere in Fig. 7 in appendix.

We have transformed these ten continuous points to six integer points corresponding to the last two columns of Table 2 in appendix, to identify cluster centers, represented successively in Figs. 8, 9 and 10 in appendix.

Finally, Fig. 11 in appendix presents differents clusters and an analytical straight line in red color that passes through all the clusters. The red line is defined by: $y = x$. That leads to $x - y = 0$ and gives $0 \leq x - y < w$ for analytical definition. That corresponds to $0 \leq x - y < 1$ for naif analytical straight line regression; $0 \leq x - y < 2$ for standard analytical straight line regression and $0 \leq x - y \leq 2$ for supercover analytical straight line regression, touching all the clusters.

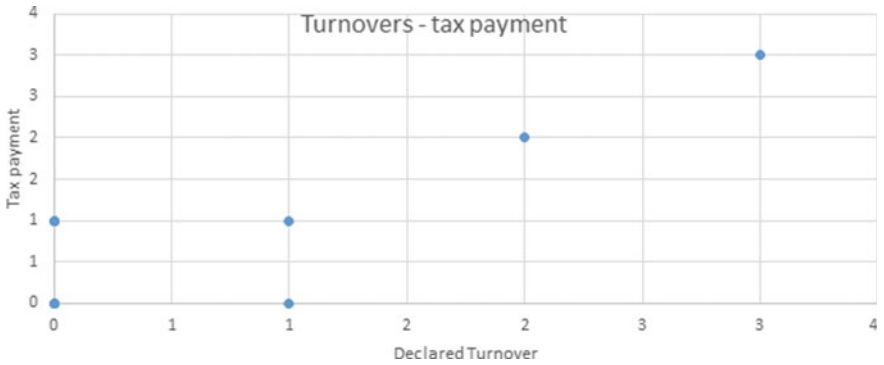


Fig. 8 Data reduction of ten continuous points to obtain six integer points, representing cluster centers, where data are from the last two column of Table 2 in appendix

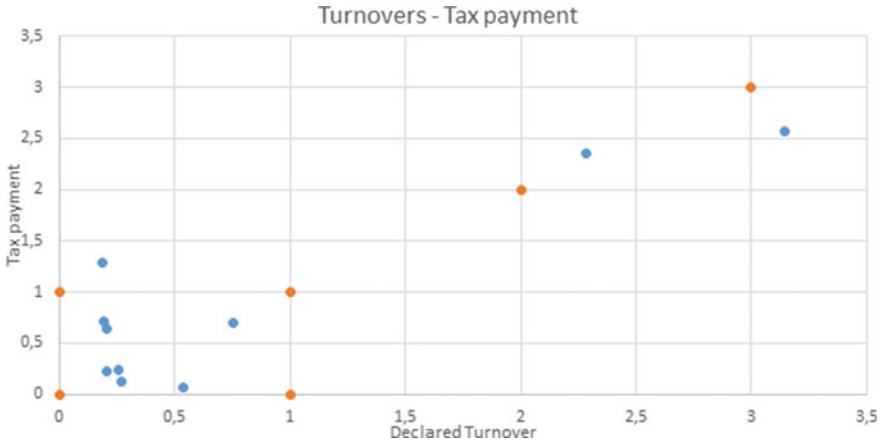


Fig. 9 Superposition of initial data on the cluster centers

The average of $|a_{j,1} - a_{i,1}|$ and $|a_{j,2} - a_{i,2}|$ will be updated with each new pair $(a_{i,1}, a_{i,2})$ and then also the values of α and β, which will contribute easily to prediction.

Moreover, instead of using the average, others functions could be used to determine α and β.

4.2 Using Arbitrary Pairs (l, K)

This section is focusing on three cases of arbitrary pairs (l, k) as parameters summarized in Table 3 in appendix, to compute the pair (α, β) as follows: $\begin{cases} \alpha = \frac{|b-a|}{k} \\ \beta = \frac{|d-c|}{l} \end{cases}$. For

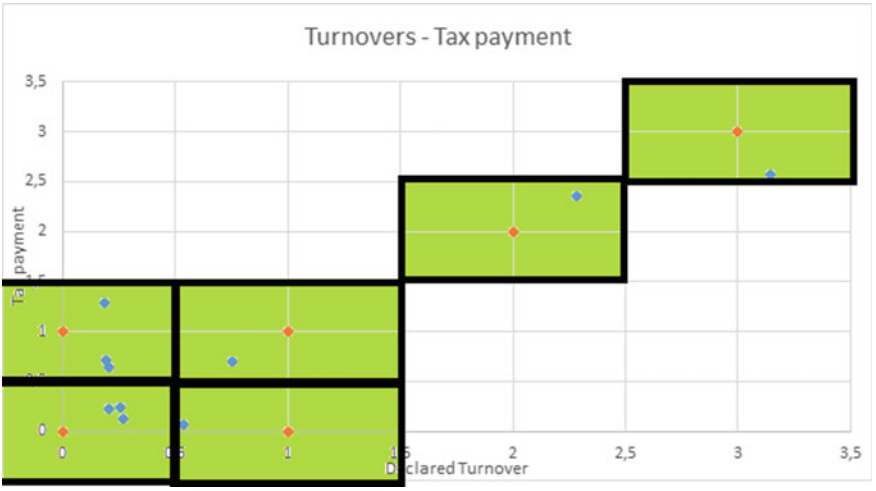


Fig. 10 Data reduction of ten continuous points to obtain six integer points, where data are from Table 2 in appendix

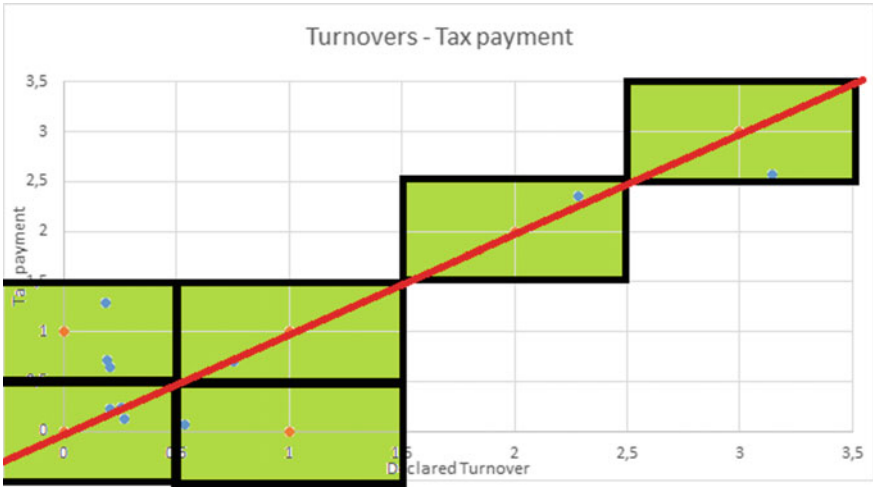


Fig. 11 Data reduction of ten continuous points to obtain six integer points, where data are from Table 2 in appendix and analytical straight line found, in red color

instance, Tables 4, 5 and 6, in appendix, show respectively the related coordinates (x, y) of the cluster center in the new repere accordingly to different values (α_0,β_0), (α_1,β_1) and (α_2,β_2) of (α,β).

For all the three cases in Table 3 in appendix, it appears several pairs in the initial repere that belongs to the same cluster in the new repere. The standard Hough transform is applied to these clusters, accordingly to the dual of their coordinates centers.

Table 3 Parameters for Data analysis in the new repere

(l, k)	α	β	(a, b, c, d)
(100, 100)	$\alpha_0 = 7629375.07$	$\beta_0 = 1219189.47$	(8.543527799, 108.5435278, 6.192568248, 106.1925682)
(1000, 1000)	$\alpha_1 = 762937.507$	$\beta_1 = 121918.947$	(85.43527799, 1085.435278, 61.92568248, 1061.925682)
(10000, 10000)	$\alpha_2 = 76293.7507$	$\beta_2 = 12191.8947$	(854.3527799, 10854.35278, 619.2568248, 10619.25682)

Table 4 Data analysis based on ten pairs (turnovers, tax payment) with $\begin{cases} \alpha_0 = 7629375.07; \\ \beta_0 = 1219189.47; \end{cases}$

Declared turnovers (x_i)	Tax payment (y_i)	Cluster x ($E(\frac{x_i}{\alpha_0})$)	Cluster $y(E(\frac{y_i}{\beta_0}))$
828,119,285	25,655,100	109	21
600,421,484	23,612,142	79	19
197,996,735	6,946,016	26	6
141,416,533	611,073	19	1
70,559,937	1,197,763	9	1
67,360,002	2,344,750	9	2
54,147,500	2,271,249	7	2
53,780,081	6,395,359	7	5
50,599,898	7,068,074	7	6
49,478,528	12,934,269	6	11

In this sense, a program in Python using OpenCV library has been implemented to compute the standard Hough transform method, applied to each rectangle (as pixel) as introduced by Cheick and others in [9], to update the accumulator data with the number of votes and to highlight analytical straight line parameter, used for linear regression.

As described in Sect. 3, the proposed technique focuses on the internal structure of datasets X_i to establish firstly clusters, before computing linear regression. Only, clusters having the number of continuous points, superior to the threshold γ will be considered in computing Hough transform. For example $\gamma = 4$, clusters having at least four continuous points (inside) are considered for dual computing in the accumulator. Due to the values of α and β, several cells having the highest vote will occur in the accumulator (or the parameter space).

Table 5 Data analysis based on ten pairs (turnovers, tax payment) with $\begin{cases} \alpha_1 = 762937.507; \\ \beta_1 = 121918.947; \end{cases}$

Declared turnovers (x_i)	Tax payment (y_i)	cluster x ($E(\frac{x_i}{\alpha_1})$)	cluster y($E(\frac{y_i}{\beta_1})$
828,119,285	25,655,100	1085	210
600,421,484	23,612,142	787	194
197,996,735	6,946,016	260	57
141,416,533	611,073	185	5
70,559,937	1,197,763	92	10
67,360,002	2,344,750	88	19
54,147,500	2,271,249	71	52
53,780,081	6,395,359	70	19
50,599,898	7,068,074	66	58
49,478,528	12,934,269	65	106

Table 6 Data analysis based on ten pairs (turnovers, tax payment) with $\begin{cases} \alpha_2 = 76293.7507 \\ \beta_2 = 12191.8947 \end{cases}$

Declared turnovers (x_i)	Tax payment (y_i)	cluster x ($E(\frac{x_i}{\alpha_2})$)	cluster y($E(\frac{y_i}{\beta_2})$
828,119,285	25,655,100	10,854	2104
600,421,484	23,612,142	7870	1937
197,996,735	6,946,016	2595	570
141,416,533	611,073	1854	50
70,559,937	1,197,763	925	98
67,360,002	2,344,750	883	192
54,147,500	2,271,249	710	186
53,780,081	6,395,359	705	525
50,599,898	7,068,074	663	580
49,478,528	12,934,269	640	1061

5 Conclusion and Perspectives

Our analysis shows that classical Hough transform is a regression technique. Exploration of Hough transform extension and its combination with tiling space are integrated as a new regression technique. In this way, linear regression of continuous points in 2D has been implemented through the combination of two techniques, precisely the application of a mesh generation followed by the standard Hough transform method in order to achieve data classification.

In experimental analyst, comparisons of this method and classical regression reveal that the method based on composition is better than classical method in the

term of tolerance of errors in dataset in the limited area of (α x β). Comparisons show effectively difference between the classical Hough transform and the proposed method based on a mesh generation and the Hough transform method: it appears best results in the term of visibility of data relation in the dataset, in playing on the size of a cluster, in varying the technique of mesh generation through α and β. Several initial data in the dataset lead to the same cluster, realizing data reduction.

In perspectives, multiple regression could be studied in n dimensional space, related to the number of variable X_i. The dual of a voxel in 3D space defined in Dexet's thesis in [7] should contribute to the application of Hough transform, to establish extension for multiple regression.

In the future works, polynomial regression can be extended to take into account analytical polynomial functions based on discrete points, as specified in the perspectives of Sere's thesis in [8]. It consists of polynomial functions, applied to clusters (square), having the number of internal continuous points, superior to a threshold γ, being initially defined.

6 Appendix

See Figs. 3, 4, 5, 6, 7, 8, 9, 10, 11 and Tables 2, 3, 4, 5, 6.

Algorithm 1: function computing(X ($a_{i,1}$), Y($a_{j,2}$) : variables of reals)

Result: computing α and β
i, j, N: integer ;
som1, som2 : real ;
som1 $\leftarrow$0; ;
som2 $\leftarrow$0; ;
N $\leftarrow C_p^2$; ;
for *i=0 to p-1* **do**
  ;
  for *j=i+1 to p* **do**
    ;
    som1 $\leftarrow$ som1+ $|a_{j,1} - a_{i,1}|$;
    som2 $\leftarrow$ som2+ $|a_{j,2} - a_{i,2}|$;
  end
end
$\alpha \leftarrow \frac{som1}{N}$;
$\beta \leftarrow \frac{som2}{N}$;

Algorithm 2: function ComputingDualofValidClusters(liste : vector of pairs of reals (x_i, y_i))

```
Result: list of pairs
d : pair of reals (x_i, y_i);
dc : pair of integer ;
c: Couple (pair, 0) ;
x, y : Real ;
t, v, cpt : Integer ;
tab : vector of pairs of integers ;
listvaleur : vector of Couple (pair, nb) ;
for each d in liste do
    x←d.getX() ;
    y←d.getY() ;
    t←E(x/α) ;
    v←E(y/β) ;
    dc← new pair(t,v) ;
    if dc not in tab then
        tab.add(dc) ;
        c←couple(dc, 1) ;
        listvaleur.add(c) ;
    end
    else
        c←couple(dc, 1);
        listvaleur.update(c);
    end
end
;
for each c in listevaleur do
    cpt← c.getNb();
    if γ <cpt then
        dc←c.getPair();
        dual(dc)
    end
end
;
return tab;
```

References

1. Bresenham JE (1965) Algorithm for computer control of a digital plotter. IBM Syst J ACM
2. Wright WE (1991) Rendering: parallelization of Bresenham's line and circle algorithms. IEEE Comput Soc Press doi 10(1109/38):59038
3. Reveillès JP (1995) Combinatorial pieces in digital lines and planes. Proc SPIE 2573, Vis Geomet IV(23)
4. Andres E (2003) Discrete linear objects in dimension n?: the standard model. Graph Models 65(1–3):92–111
5. Sere A, Traore Y, Ouedraogo FT (2018) Towards new analytical straight line definitions and detection with the Hough transform method. Int J Eng Trends Technol 62(2):66–73
6. Hough P-V-C (1962) Method and means for recognizing complex patterns. In United States Pattent 3069654:47–64

7. Dexet M (2006) Architecture d'un modeleur géométrique à base topologique d'objets discrets et méthodes de reconstruction en dimensions 2 et 3, Université de Poitiers (France) Thèse en informatique et applications
8. Abdoulaye SERE (2013) Transformations analytiques appliquées aux images multi-échelles et bruitées, thèse de doctorat unique en informatique, Université de Ouagadougou
9. Traore CADG, Sere A, Straight line detection with the Hough transform method based on a rectangular grid. In: Fifth international conference on information and communication technology for competitive strategies (ICTCS-2020)
10. Duda RO, Hart PE (1972) Use of the Hough transform to detect lines and curves in pictures. Communications of the ACM 15(1):11–15
11. Fanelli G, Gall J, Goo LV (2009) Hough transform-based mouth localization for audio-visual speech recognition. In: British machine vision conference
12. Yao A, Gall J, Van Goo L (2010) A Hough transform-based voting framework for action recognition. In: IEEE conference on computer vision and pattern recognition (CVPR 10)
13. Dexet M, Andres E (2009) A generalized preimage for the digital analytical hyperplane recognition. Discrete Appl Mathematcs 157(3):476–489
14. Sere A, Sie O, Andres E (2012) Extended standard Hough transform for analytical line recognition. In: 6th International conference on sciences of electronics, technologies of information and telecommunications (SETIT), Sousse, Tunisia, pp 412–422. https://doi.org/10.1109/SETIT.2012.6481950
15. Sere A, Traore CADG, Traore Y, Sie O (2020) Towards traffic saturation detection based on the Hough transform method. In: Proceedings of the future technologies conference (FTC) 2020, Volume 2. FTC 2020. Advances in Intelligent Systems and Computing, vol 1289. Springer, Cham
16. Spratling MW (2016) A neural implementation of the Hough transform and the advantages of explaining away. Image Vis Comput. https://doi.org/10.1016/j.imavis.2016.05.001
17. Bamogo M (2021) application de la transformée de Hough et les réseaux de neurones dans la détection des droites discrètes, mémoire de master en informatique, Ecole Supérieure d'Informatique (ESI), Université Nazi BONI

Data Acquisition Techniques from IOT Devices for Smart Transportation: A Brief Overview

V. Ranjith and Kiran B. Malagi

Abstract Internet of Things has revolutionized the entire world these days. With the advent of the word smart in all the fields, for example, the smart city, huge amount of data gets accumulated every single day. For any application, processing this huge amount of data is very important. Prior to processing, acquisition of this bulk amount of data is very important. In the existing research, there are different ways to collect these data. In this paper, data acquisition using different modern approaches is outlined. These approaches encompass data acquisition using hierarchical deep reinforcement learning, data acquisition using energy efficient UAV, data acquisition using swarm intelligence. This paper outlines these approaches, brings out the assets and liabilities of each one of these approaches for smart com.

Keywords UAV · Deep reinforcement learning · Blockchain · SADOL · MADOL

1 Introduction

Smart city has been one of the major applications of Internet of Things. The Internet of Things has transformed the entire world spruce. Smart transportation is one of the major applications of Internet of Things. If we take up any smart transportation technology, for example, if there is a smart car and the car may have to travel from a specific source to destination and let assume that there are many routes from that source to destination. In such cases, it is required to find the optimal route from that source to destination. Henceforth, we rely upon the data that need to be collected from things in the Internet. The things in the Internet may be sensor nodes, cameras, and many more devices. Therefore, lots of data need to be collected from such remote devices.

V. Ranjith (✉) · K. B. Malagi
Dayananda Sagar University, Bangalore, India
e-mail: leoxranj@gmail.com

K. B. Malagi
e-mail: kiranmalagi-cse@dsu.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_6

There are many traditional methods of data collection for smart transportation [1]. One of such methods is by using curvaceous tubes. These curvaceous tubes are installed on roads. When the vehicles pass on these tubes, cram air would be released from the tube. For each of these vehicles that have passed the tubes, the resulting cram sends a burst of air to air switch. These air switches are installed on traffic calculating device [2]. The cram sends the csv format of data collected to the traffic calculating device. The merit associated with this technique is that it is malleable. The demerit associated with this technique is that it is traffic reliant. The second traditional technique for collecting the IoT data is inferential loops. The inferential loop consists of embedded twisted cord, waver, and loop extension cord. Whenever the vehicles pass over the loop, the corresponding count would be recorded in the calculating device. The resulting output of this technique is the computed value of number of vehicles that have run over the loop. The merit associated with this technique is that it is not expensive and easily maintainable. The demerit associated with this technique is that it requires more than one loop to take the calculation of number of vehicles that have passed over. The third method of data collection is video conveyance detection. This is the secretive and discreet technique. This technique works as follows. Firstly, it computes the number of vehicles that are in the front. Secondly, it computes the number of vehicles that are in the back. Finally, it subtracts the number of vehicles that are in the back by number of vehicles that are in the front. The merit that is associated with this technique is that it is extensible [3]. The demerit that is associated with this technique is that it is expensive to maintain. The next technique is based on global positioning system. In this technique, the global positioning systems that are installed in the car collect the speed and other data from satellites. It collects the precise and error-free data [4]. The merit associated with this technique is that it is cheaper compared to other techniques. The next technique is by using fog sensor. As we know, smart transportation is not only about finding the optimal route. It is all about looking on to other factors for safe traveling. So, sometimes, especially during winter season, there will be lot of fog that gets accumulated on the vehicles. So, cleaning up of this fog is very important for safe transportation of vehicle. Therefore, the concept of fog sensor was introduced. With this technique, the fog can be monitored in a contemporized and discreet way [5]. The next technique to find out the traffic mass is by using ultrasonic sensors. This technique works based on the binary digits. The merit associated with this technique is that it is error free and easily transferable. It is useful in finding the traffic mass and discreet technique having high error-free feature. The next technique is by using magnetometer sensor. This technique collects the data for vehicle speed. It is used as an alternative to inferential loops. The merit associated with this technique is that it is coincided and discreet. The abovementioned techniques are few of the traditional methods that can be used for data collection or data acquisition. Nowadays, with the advent of machine learning models, data acquisition has become more reliable and efficient. In this paper, we outline few data acquisition techniques using machine learning.

2 Modern Techniques

2.1 UAV-Aided Data Acquisition Using Swarm Intelligence

There are huge number of devices in the Internet of Things. The vivacity that these devices are having is very less [6]. There are situations where a greater number of IoT devices will be located in diverse demographic locations. The devices may be located somewhere near the oceans or located somewhere near the mountains. In such cases, these devices consume lot of vivacity which is not an efficient case. Secondly, the IoT devices may be very much influenced by the environment which leads to greater loss of vivacity. Therefore, the concept of unmanned aerial vehicles was introduced. The UAV acts as a communicating agent between the base station and the IoT devices. The UAV carries its own vivacity, so whenever it is flying from the base station to the IOT device, it can recharge itself recursively. Also, if any IOT device is located in the ocean beds, it can communicate with one more device that is located on the water surface of the ocean. This device acts as a mediator between UAV and the device that is located in the water. Therefore, this is how the UAVs help in energy efficient operation of data collection [7]. The regions where the IoT devices are located are divided into different divisions. Further, the divisions where the UAVs have to collect data are made specific [8]. For this purpose, the concept of swarm intelligence was introduced. Swarm intelligence is mainly used for disposition of UAVs. The data collection from IoT devices with the help of UAVs is shown in Fig. 1

The purpose of UAV disposition escalation is to reduce the vivacity consumption of both IoT devices and UAV. There are other techniques for disposition of UAVs. This technique works as follows. Firstly, the land points are predetermined. Later on, they are processed one after the other. For this, we use DEVIPS algorithm. By using this algorithm, the location of land algorithm gets adjusted automatically at the time of data collection. This method is very straightforward and uncomplicated, but its search is inefficacious. For this reason, swarm intelligence was introduced which identifies the landing point as a function of three coordinates. The number of landing points can be inserted or discarded arbitrarily. The demerit associated with this technique is that it does not address the elevation of the UAV. Also, the landing points cannot be interactively found every time.

2.2 Data Acquisition Using Deep Reinforcement Learning in UAVs

In the case of Internet of Things, large number of nodes will be connected. These nodes may be rechargeable or battery operated. To overcome this drawback, the disperse node can be used. These disperse nodes may not be required to recharge or battery operated, but the drawback associated with these disperse nodes is that the conveyance range of these devices may be not in boundary. To address this challenge,

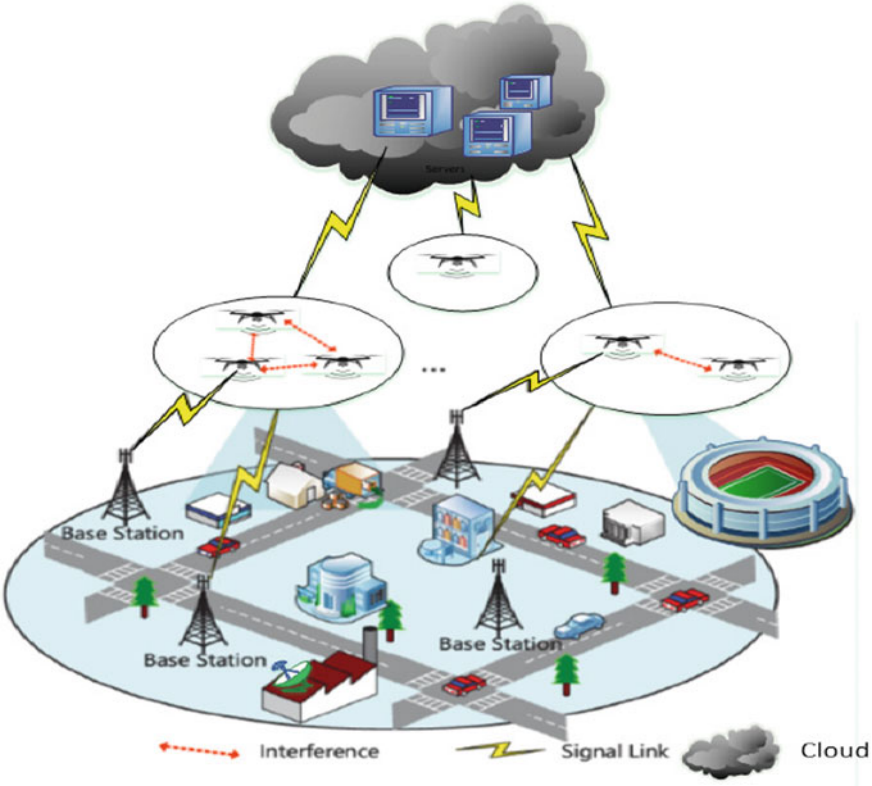


Fig. 1 UAV collecting data from IoT devices

UAVs [9] were introduced. In this technique, the UAVs can reach to the disperse nodes and collect the data individually. At the time of data acquisition, the UAVs can come back and recharge themselves again. In this technique, the UAV initiates itself from the rechargeable station and moves to the disperse nodes network. These disperse nodes can be charged by RF signals carried by UAVs. The UAVs initiates themselves and issues an inquiry signal to the disperse nodes. These disperse nodes update the UAVs with the data. Hence, the data collection is achieved. The procedure can be summarized as follows:

i. The Gaussian mixture model is used to divide the disperse nodes into bundles, and the UAVs would process each bundle individually. The bundles may be inescapable and cryptic.
ii. For inescapable bundles, we use single-agent deep option learning algorithm, and for cryptic bundles, we use multi-agent deep option learning algorithm.
iii. In this technique, the MADOL algorithm is scattered. Figure 2 can be shown.

Fig. 2 UAVs using disperse nodes

The merit associated with this technique is that it is efficient than other RL techniques. The demerit associated with this technique is SADOL algorithm takes more time.

2.3 Data Acquisition Using Deep Reinforcement Learning for Multi-UAVs

In Internet of Things, lot of devices will be connected in the network. The data that are being collected by the IoT devices will be sent to the central coordination center [10]. Since both UAVs and IoT devices consume lot of vivacity when they are operating, there is a need to reduce the vivacity consumption of both of these devices. Also, the UAVs can easily readjust their paths if they come across any obstructions to reach the IoT device [11]. The steps in this process can be summarized as follows:

i. **UAV map-reading**: In this step, the UAV finds the IoT device from where the data have to be collected by overcoming the obstructions if any. This process is done by using the scope determiners.
ii. **Device current dominance**: In this step, the UAVs collect the data from the IoT devices. The UAV sets the clear transmit power for the IoT device. In this way, the UAVs achieve the vivacity efficacy of the IOT devices.
iii. **Multiple UAV Organization**: In this technique, the data collection is planned in a proper way. It is made sure that the UAVs do not run out of vivacity.

For all the above processes, we use deep reinforcement learning algorithms. The merit associated with this technique is that it achieves efficiency in terms of vivacity. It provides better solution. The demerit is that it is more complex.

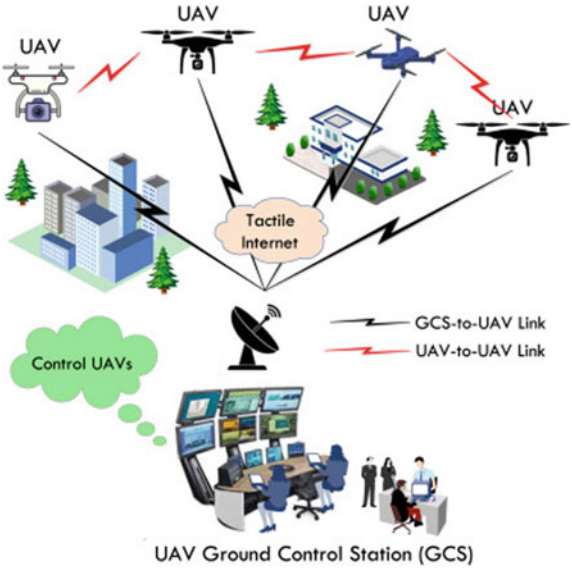


Fig. 3 Blockchain-enabled UAVs

2.4 Data Acquisition Using Blockchain in UAV

In the normal process, UAVs may run out of their vivacity, and since the IoT devices are externally visible to the environment, these devices may be subjected to attacks. To overcome these problems, blockchain technology was added to the existing process. By pushing data and transactions, UAVs get charging coins as benefits. To provide security, this technique uses blockchain to protect the IoT devices from unwanted invasions. The merit associated with this technique is that it uses prediction algorithm to reduce vivacity and provide protection for IoT devices. The demerit associated with this technique is that it is more complex. Figure 3 for the above process is shown.

The table below summarizes the above techniques.

Methods	Merits	Demerits
Data acquisition using swarm intelligence	Less complex	Elevation of UAV not addressed
Deep reinforcement learning	Provides better efficiency	SADOL algorithm is not time efficient
Deep reinforcement learning using multiple UAVs	Efficient vivacity	More complex
Blockchain-enabled UAV	Reduces vivacity consumption using prediction	More complex

3 Conclusion

In this paper, we have identified and looked over the various modern data acquisition techniques. We have outlined each and every techniques. We have also identified the merits and demerits of these techniques. Though these techniques are having merits, they are some way not better than each other. Based on their merits and demerits, we can select a particular technique for data collection from IoT devices for smart transportation. In the next paper, we would outline the different methods of providing security for IoT data.

References

1. Bawany NZ (20115) Smart city architecture: vision and challenges 6(11):246–255
2. Leduc G (2008) Road traffic data: collection methods and applications
3. Chintalacheruvu N, Muthukumar V (2012) Video based vehicle detection and its application in intelligent transportation systems 2012(October):305–313
4. Rathore MM, Ahmad A, Paul A, Jeon G (2015) Efficient graph-oriented smart transportation using internet of things generated big data. In: 11th International conference signal-image technology internet-based system, pp 512–519
5. Crosby JD, Visibility sensor accuracy: what's realistic, pp 1–5
6. Arabi S, Sabir E, Elbiaze H, Sadik M (2018) Data gathering and energy transfer dilemma in UAV-assisted saying access network for IoT. Sensors 18(5):1519
7. Mozaffari M, Saad W, Bennis M, Debbah M (2016) Mobile internet of things: can UAVs provide an energy-efficient mobile architecture? In: Proceeding IEEE global communication conference (GLOBECOM). Washington, DC, USA, pp 1–6
8. Huang P, Wang Y, Wang K, Yang K, Differential evolution with a variable population size for deployment optimization in a UAV-assisted IoT data collection system. In: IEEE transaction emerging topics computer intelligent, vol 4. https://doi.org/10.1109/TETCI.2019.2939373
9. Zhang S, Zhang H, Di B, Song L (2019) Cellular cooperative unmanned aerial vehicle networks with sense-and-send protocol. IEEE IOT J 6(2):1754–1767
10. Ghdiri O, et al (2020) Energy-efficient multi-uav data collection for IoT networks with time deadlines. In: Proceeding IEEE global communication conference

Analysis of Student Behavioural Patterns by Machine Learning

L. V. Krishna Rao, B. V. Gowthami, B. Hema, A. Sai Saketh, and G. Narendra Babu

Abstract An important task in education field is discovering student behavioural patterns to take timely action to improve student activities or grades. Sometime students may fell into depression due to misunderstanding of subjects or due to low grade which leads into abnormal behaviour, and by identifying such abnormal behaviour, institutions can take necessary steps to improve student's condition. For this research, questionnaire method is used which includes collecting student data through survey and analyse students' behavioural patterns. However, results by this method are not effective or accurate as this method largely relies on feedback data. So to solve this problem, an unsupervised clustering approach can be used. This produces relatively accurate results. The proposed framework integrates two unsupervised clustering approaches, i.e. density-based spatial clustering of applications with noise (DBSCAN) and *k*-means. The students data is collected from Kaggle data sets. The proposed framework extracts necessary behaviour features by statistics and entropy to find both anomalous behavioural patterns and main stream patterns. To predict whether the student is low active or high active or medium active, we can use supervised techniques as unsupervised clustering approaches are meant to form clusters. These findings can help students to improve their grades and personality and organization can also take appropriate steps to help students by providing better services and administrations such as psychological consultations and academic advices.

Keywords Ensemble clustering · DB SCAN · *K*-means · Supervised techniques

L. V. Krishna Rao · B. V. Gowthami (✉) · B. Hema · A. Sai Saketh · G. Narendra Babu
Lakireddy Balireddy College of Engineering, NTR Dt, Mylavaram, AP 521230, India
e-mail: gowthamibevara26@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_7

1 Introduction

Educational organizations should understand behavioural patterns of students and should take best steps to improve students' behaviour. For example, finding correlation between student behaviours and their academic performance [1–4]. First, correlation between sleeping habits and academic performance. According to the recent research by Arne H. Eliasson, Christopher J. Lettier and Arn. H. Eliason, student performance is effected by total sleep time of students. A questionnaire is conducted in October 2007 which includes questions related to study habits, number of sleeping hours, academic performance and reasons for wakefulness. As it is observed that there is no significance difference in total time of sleep of both low-graded and high-graded students, it is concluded that sleeping times and wakefulness are more closely correlated to academic performance of students than total time of sleep of students. Second, correlation between eating behaviour and academic performance. This research was performed by M. Valladares, E. Duran and colleagues in Chile. These requested their students to fill the survey which provides them data about students' eating behaviour and academic performance. Three-factor questionnaire is used to evaluate eating behaviour, and grade point average (GPA) is used to measure academic performance. They concluded that there is a positive correlation between women's eating behaviour and academic performance than men. Third, students' course performance in a massive open online course (MOOC). MOOCs in curriculum are developed by a team of doctoral students with mentorship from two professors. They offered two MOOCs, powerful tools for teaching and learning: Digital story telling MOOC and powerful tools for teaching and learning: Web 2. 0 tools on Coursera platform. This paper reports that students whose engagement is actively high in MOOCs performed better in their academics than the students who did not participate. Lastly, correlation between procrastination behaviour of students and their academic performance. Research on this was done by D. Hooshyar, M. Pedaste and Y. Yang. They used linear support vector machine classifier to classify the students based on their amount of procrastination. It is concluded that the students who procrastinate more had low academic performance than the students who procrastinates their work less.

Most researchers used a questionnaire method to complete these studies (correlation between students' behaviour patterns and their academic performances). As a part of questionnaire, they collected specific students data from specific universities with specific features. However, this method of collecting student data has some consequences which may lead to false predictions of students category identification. First, as surveys are conducted on a scheduled day, i.e. either one per academic year or semester, it is not possible to collect students' current data in timely manner. If the anomalous behavioural patterns of students are not detected timely, it may lead to serious consequences [5–7] such as early dropouts of students. Second, false information can be provided during survey like students with anomalous behaviour may give normal answers to the questions in survey which make them appear as mainstream students and mainstream students may not fill the survey carefully due to their personal problems or any other reasons. This leads to false conclusions and

bias the analysis results. Third, to analyse students' behavioural patterns, rich expert knowledge is required to prepare questionnaire that can collect enough and correct information about student's behaviour. These consequences make the questionnaire method unreliable and uneffective.

Machine learning techniques can give accurate results than questionnaire method. Machine learning approaches are categorized as supervised, semi-supervised and unsupervised methods. Supervised learning is an approach, where an input data that has been labelled for particular output is trained. Model is trained until relationships between input data and output labels are detected to determine which class a never-before-seen student belongs to. Semi-supervised learning is an approach in which training data consists of small amount of labelled data and large amounts of unlabelled data. Unsupervised learning is used to clusters the data of students. As students behaviour keeps evolving, model must be updated timely which may not be possible using supervised and semi-supervised so unsupervised approaches are used widely in practical applications.

The rest of this paper is organized as follows: in Sect. 2, we discussed the proposed clustering framework. We then describe the algorithms we used in the proposed work in Sect. 3. Next, we describe the experimental results Sect. 4. Finally, we conclude our work and propose future work in Sect. 5.

2 Framework of Study

See Fig. 1.

2.1 Data Set Collection

Data set we used here is taken from Kaggle website. We used xAPI-Edu-Data. csv. This data set consists of 480 samples and 17 features. Gender, Nationality, Place of Birth, stageId, GradeId, SectionId, Topic, Semester, Relation, raised hands, VisITed Resources, Announcements View, Discussion, Parent Answering Survey,

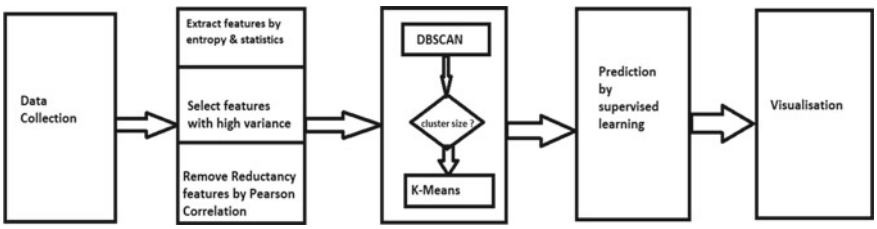


Fig. 1 Framework

Fig. 2 Screen shot of code which is displaying features in taken data set

Parent school Satisfaction, Student Absence Days, Class. Below screen is showing records of student performance data set (Fig. 2).

2.2 Feature Extraction

In above data set, some columns contain non-numeric values and clustering algorithm will not take non-numeric values so we will be applying reprocessing technique to convert non-numeric values to numeric by replacing MALE with 0 and FEMALE with 1 in gender column. Data set contains 17 columns and all columns are not important so we apply principal component analysis (PCA) algorithm which calculate importance of each features and select only those important features, and by applying this algorithm, we are selecting best 4 features from data set. Those 4 features are raised hands, VisITed Resources, Announcements, Discussions.

As shown in framework in this proposed work, we used statistics, entropy, variance and Pearson correlation to extract features. Statistics are categorized into central tendency which includes mean, mode, median and dispersion which includes variance and standard deviation. Entropy is used to measure the randomness in the information provided in data set. Entropy can be calculated by the formula.

$$H = -\sum_i p(i)\log(p(i))$$

where $p(i)$ indicates probability of ith behaviour event. Higher the entropy, it is difficult to draw conclusions from the provided data, for example: flipping a coin. Select high variance features and remove reductant, small variance features. To find how two features are correlated, we used Pearson correlation coefficient. Value of Pearson correlation coefficient is generally represented by 'r'. value of 'r' ranges from -1 to 1. If $r = 0$, it indicates there is no relation between those features. If $r > 0$, then the two features have positive relationship. If $r < 0$, they have negative relationship, i.e. their values are inversely proportional (Fig. 3).

Fig. 3 Graph showing the behaviour of Pearson correlation

2.3 Clustering Analysis

Density-based spatial clustering of applications with noise (DBSCAN) [8] and *k*-means are two unsupervised approaches which are used efficiently in daily applications. In initial step, DBSCAN is applied on data set to get clusters. Anomalous behaviour students go to noise clusters, and main stream students go to normal clusters. As the mainstream students cluster size is large, they may further contain anomalous behaviour students; so by applying *k*-means clustering, the single normal cluster can be reduced to small sized multiple clusters. The cluster with least size is considered as anomalous behavioural students records.

2.4 Prediction

To predict the class of never-before-seen student, we trained the data by using supervised approach. This takes four features, namely raised hands, VisITed Resources, Announcements, Discussions and predict the class of student, i.e. whether that particular student belong to low active or medium active or high active. As accuracy of random forest algorithm is more when compared to other supervised algorithms, random forest algorithm is used to predict the never-before-seen student's class (Fig. 4).

2.5 Visualization

The clusters formed by DBSCAN and *k*-means are visualized so that anyone can easily differentiate the students class.

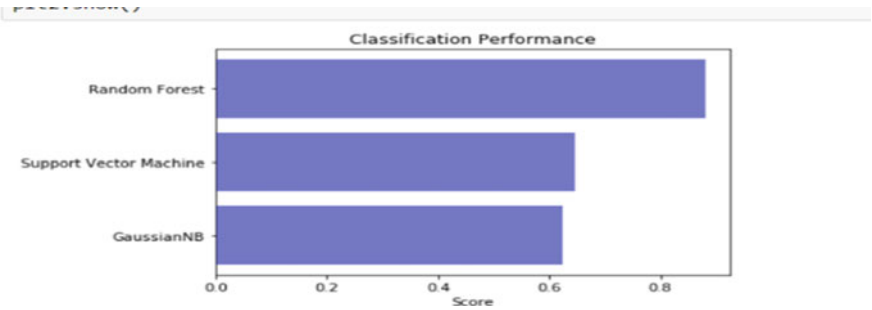


Fig. 4 Graph showing accuracy of various supervised algorithms

3 Algorithms

3.1 Initial Clustering Using Density-Based Spatial Clustering of Application with Noise(DBSCAN)

DBSCAN is one of the unsupervised approaches that is used widely in daily applications. Unlike others, it deals with noise data. Irrelevant data that effects data analysis significantly is termed as noise data. DBSCAN algorithm automatically filters noise and outliers. It requires two parameters eps and minPts. Eps defines the neighbourhood of a given radius, and minPts defines minimum number of samples that every cluster should contain. MinPts must be greater than or equal to number of dimensions in data set to get accurate results. Based on the values of minPts and eps, the data points are categorized into three types core, border and noise.

3.2 Subdivision Clustering Using K*-Means*

K-means is used to subdivide the data set into *k* number of clusters. So it acts as complement to DBSCAN. To determine the value of *k,* the elbow method is used. Elbow method uses within cluster sum of square (wcss) for each *k* value. Wcss is also known as k_inertia. Plots a graph between wcss and *k*-means, the bend point which looks like arm is considered as optimal value of *k* (Fig. 5).

3.3 Random Forest Algorithm

Random forest is popular supervised approach which used effectively to predict the class of never-before-seen data. Rather than relying on single decision tree, it takes

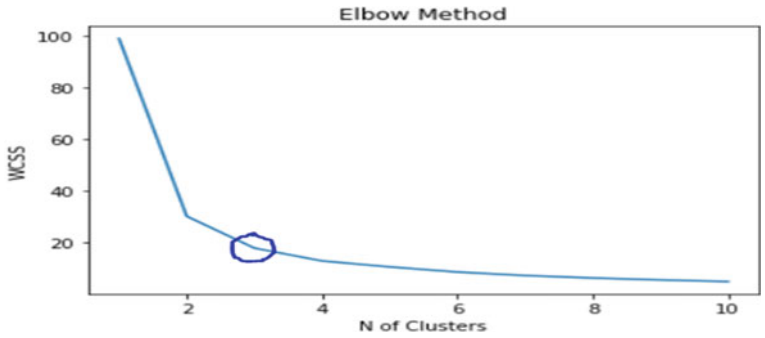


Fig. 5 Elbow method graph to calculate accurate K value

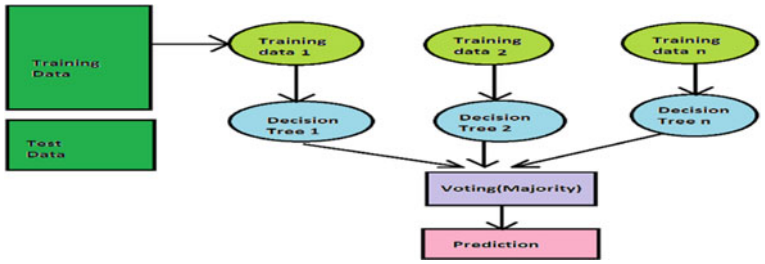


Fig. 6 Random forest algorithm

prediction from multiple decision trees, based on majority votes it predicts the final output (Fig. 6).

4 Experimental Results

We used Python language to implement machine learning models. We followed steps given below.

(1) We uploaded data set which is available in Kaggle (Fig. 7).

(2) Feature extraction is done by using entropy, Pearson correlation coefficient and variance.

(3) Clustering analysis (Fig. 8).

As we discussed earlier about elbow method by which we got optimal value of k as 3 so by applying k-means 3, clusters are formed. WCSS (inertia) is the distance between data point and centroid (Fig. 9).

```
In [5]: import pandas as pd
        import seaborn as sns
        import numpy as np
        import matplotlib.pyplot as plt

        C:\Users\Lenovo\Anaconda3\lib\site-packages\statsmodels\tools\_testing.py:19
        Use the functions in the public API at pandas.testing instead.
          import pandas.util.testing as tm

In [6]: dataset = pd.read_csv('data.csv')

        dataset.insert(loc=0, column='row_num', value=np.arange(len(dataset)))
        dataset.head(5)

Out[6]:
```

	row_num	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic	Semester
0	0	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F
1	1	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F
2	2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F
3	3	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F
4	4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F

```
In [7]: dataset.info()

        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 480 entries, 0 to 479
        Data columns (total 18 columns):
```

Fig. 7 Screenshot of code in which data set is uploading

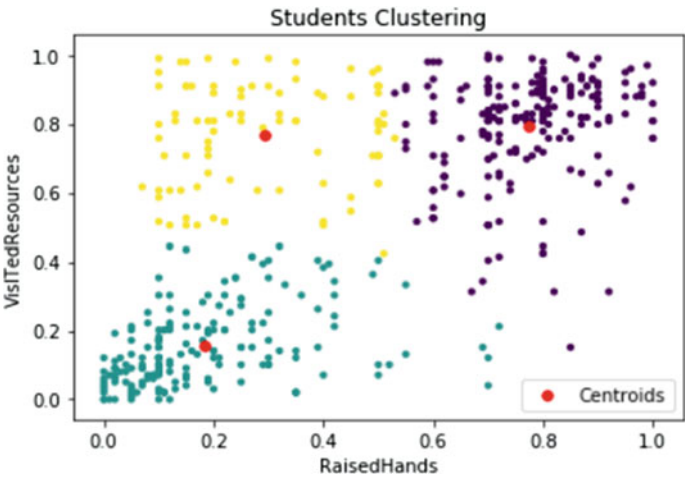


Fig. 8 *K*-means graph

The above is the visualization of DBSCAN. Black coloured data points are the anomalous students records

(4) Prediction (Fig. 10).

By taking the student details, it provides us the information about the student's class whether he belong to low active or high active or medium active class.

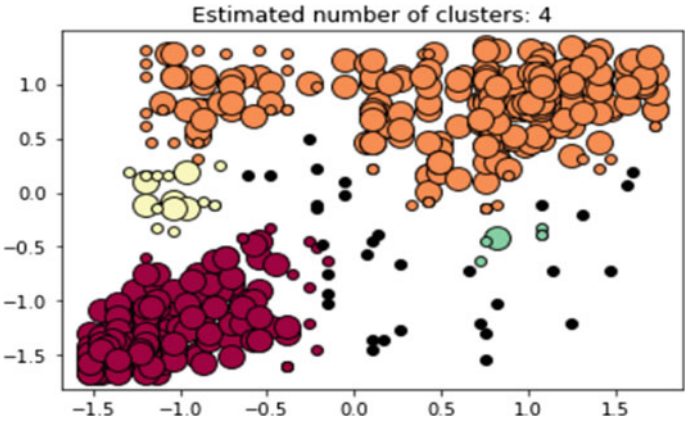


Fig. 9 DBSCAN graph

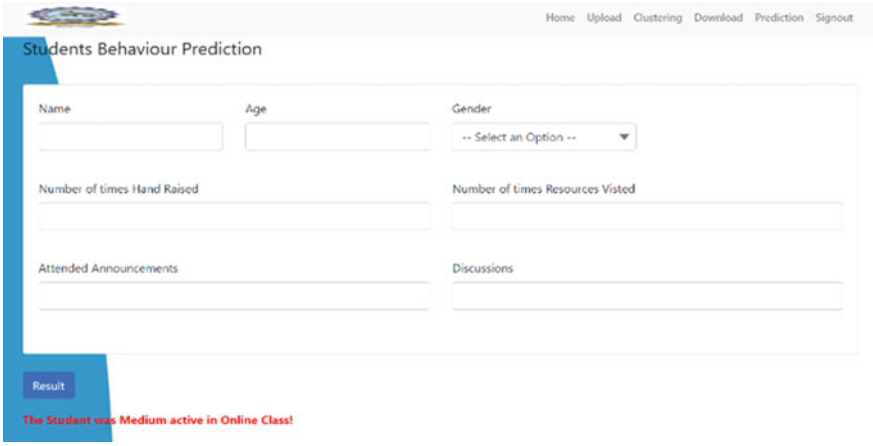


Fig. 10 Prediction screen

5 Conclusions

This paper proposed an ensemble unsupervised approach for analysis of students' behaviour patterns and a supervised approach to predict the class of never-before-seen student's class. We extracted four features from the data set we collected from Kaggle. This proposed work can detect both anomalous and mainstream students records. Based on the clustering analysis results, the educational organizations and management can take necessary measure to improve students grade and personality. For better clustering analysis, future work should include the following (1) Extract more meaningful multisource behavioural data. (2) For high-dimensional feature spaces design, a new distance measures to make proposed work more effective. (3)

Study the relationship between students' academic performance, psychological state and employment domain.

References

1. Eliasson AH, Lettieri CJ, Eliasson AH (2010) Early to bed, early to rise! Sleep habits and academic performance in college students. Sleep Breathing 14(1):71–75. https://doi.org/10.1007/s11325-009-0282-2
2. Valladares M, Duran E, Matheus A, Duran-Agueero S, Obregon AM, Ramirez-Tagle R (2016) Association between eating behavior and academic performance in university students. J Amer College Nutrition 35(8):699–703. https://doi.org/10.1080/07315724.2016.1157526
3. Phan T, McNeil SG, Robin BR (2016) 'Students' patterns of engagement and course performance in a massive open online course. Comput Edu 95:36–44. https://doi.org/10.1016/j.compedu.2015.11.015
4. Hooshyar D, Pedaste M, Yang Y (2019) 'Mining educational data to predict Students' performance through procrastination behavior. Entropy 22(1):12. https://doi.org/10.3390/e22010012
5. Iam-On N, Boongoen T (2017) Improved student dropout prediction in Thai university using ensemble of mixed-type data clusterings. Int J Mach Learn Cybern 8(2):497–510. https://doi.org/10.1007/s13042-015-0341-x
6. Harwati I, Virdyanawaty R, Mansur A (2015) Drop out estimation students based on the study period: comparison between naive Bayes and support vector machines algorithm methods. In: Proceeding ICET4SD. Yogyakarta, IN, USA
7. Aparicio-Chueca P, Maestro-Yarza I, Domínguez-Amorós M (2016) Academic profile of students who drop out a degree. A case study of faculty of economics and business, UB. In: Proceeding EDULEARN. Barcelona, Spain
8. Ester M, Kriegel H, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceeding 2nd international conference knowledge discover data mining, pp 226–231

Similar Incident Detection for IT Service Management (ITSM) in Multilingual Records Using Information Retrieval Model

S. B. Rajeshwari and Jagadish S. Kallimani

Abstract Integrated IT service management (ITSM) follows ITIL service operations process to track service request, incident, problem, or change in existing application or infrastructure stack. These tickets are raised by end-user, developer, and tester or through proactive monitoring sources to report a problem or request new feature development. This work contributes to AIOps offering for site reliability engineer (SRE) team who needs to continuously improve the application availability, security, and performance by analyzing re-occurring issues, frequent changes, performance metrics, etc. This solution also has feedback mechanism which continuously trains the model on result relevance. This research work will help the SRE to identify re-occurring issues irrespective of language, accelerates root cause investigation using historical resolution details and permanently fix the problem hampering the application performance. It is noticed that 80% of the resolution time is spent in diagnosing and detecting the issue. With knowledge of historical similar issues and resolution information, the time to detect and troubleshoot reduces drastically from 80 to 20%.

Keywords ITSM · SRE · ITIL · AIOps · Watson translator · Multilingual records · Information retrieval · IT service operation framework · MTTR · DevOps

1 Introduction

Due to the multinational user base, most of the user generated tickets follow a multilingual text which gets stored in native language, which makes it difficult for an

S. B. Rajeshwari · J. S. Kallimani (✉)
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore and Affiliated to Visvesvaraya Technological University, Belagavi, India
e-mail: jagadish.k@msrit.edu

S. B. Rajeshwari
e-mail: rajeshwari.sb@msrit.edu

Department of Artificial Intelligence and Machine Learning, M S Ramaiah Institute of Technology, Bangalore, Karnataka, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_8

SRE to analyze re-occurring issues, common requests, etc. Along with the lingual problem, mining such a large text data and tagging similar and common problem patterns manually is a tedious task which can easily take more than a month. As part of this work, building a similar incident retrieval model for multilingual records is considered. It parses all historical issues, detects language through business service geography, and translates non-English languages to English using IBM Watson translator. In order to analyze frequent issues, SRE needs to go through all the incidents reported by end-user as well as events alerted by monitoring systems.

In the twenty-first century, IT infrastructure and applications act as the central nervous system for businesses, governments, and other complex organizations. In today's technology-driven landscape, even the smallest of IT outages can cause massive economic impact [1]. IT leaders indicate that the costliest aspects of downtime are lost revenue (53%), lost productivity (47%), and negative impact on corporate reputation (41%).

IT teams shift through data pulled from topology, logs, tickets, alerts, and more to better predict and reactively solve IT outages. IT professionals around the world are overwhelmed with data but often have difficulty identifying the insights that truly matter. SRE can spend the majority of their time sifting through multiple data sources, but they need to shift to mission-critical workloads. Teams must be agile to operate across multiple departments quickly [2].

2 Background—ITIL Service Operations Framework

The objective of ITIL service operations is to make sure that IT services are delivered effectively and efficiently by fulfilling user requests, resolving service failures, fixing problems, and carrying out routine operational tasks. Incident management manages the lifecycle of all incidents to return the IT service to users as quickly as possible [3]. It provides permanent or temporary resolution to end-users to restore normal operations. Problem management manages the lifecycle of problems to prevent incidents from happening, and to minimize the impact of incidents that cannot be prevented [4]. Root cause analysis (RCA) is performed to find a permanent solution. Facilities management manages the physical environment where the IT infrastructure is located, including things like power and cooling, building access management, and environmental monitoring [5]. Technical management provides technical expertise and support for the management of the IT infrastructure. Closure set up an automation rule to automate ticket closure process or let end-users close tickets them through the self-service portal [6].

Service operations oversee the daily overall health of the services that are delivered to users and business functions. This includes managing disruptions to service and ensuring restoration after incidents and determining the root cause of problems [7]. Figure 1 shows the general work flow of service management. It shows the event management tool along with an incident reported.

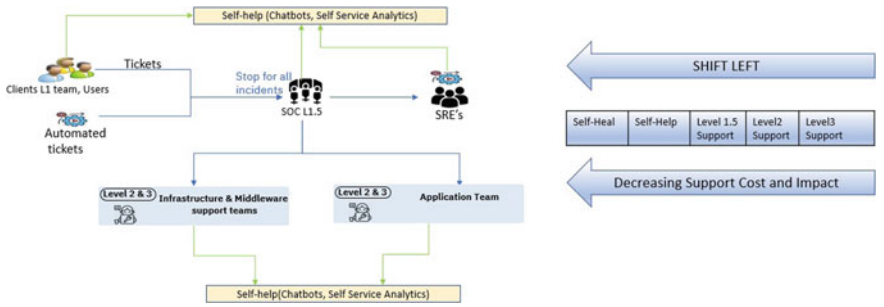


Fig. 1 Service management through level 1.5, SRE's, self-service, and automation

3 Problem Statement

SRE team needs to continuously improve the application availability, security, and performance by analyzing re-occurring issues, frequent changes, performance metrics, etc. In order to analyze frequent issues, SRE needs to go through all the incidents reported by end-user as well as events alerted by monitoring systems along with system logs, all at the same time. This is quite overwhelming to connect the dots and investigate the root cause. Due to the multinational user base, most of the user generated tickets follow multilingual texts. It gets stored in native language. It is seen that 80% of the resolution time is spent in diagnosing and detecting the issue. With the knowledge of historical similar issues and resolution information, the time to detect and troubleshoot reduces drastically from 80 to 20%.

Even though similar service is available in elastic search, however, this work is unique due to its key features like.

- The uniqueness of the similar incident detection is that it can parse multilingual records and retrieve relevant records as soon as a new incident is reported.
- The solution has feedback mechanism which takes real-time feedback from the multiple resolvers as binary input and stores it for future reference.
- The microservice built can be used as a re-usable asset for any project with same input features, i.e., issue text description.
- This microservice can be called via API and serve through collaboration platform, ticketing tools, or stand-alone curl commands.
- The output received can be directly updated in the ticket resolution details to further enrich the solution details to be used in advanced analytics for continual improvement.
- It is easily deployable as a microservice in any environment in a lightweight and portable manner as app containers on docker.

4 Objectives

The main objective of this work is to build a similar incident retrieval model for multilingual records which meets below functional requirements.

- Parses all historical issues, detects language through business service geography, and language selected in the ticket.
- Calls a Watson translator REST API if non-English language is detected.
- Translates non-English languages to English using IBM Watson translator and stores the results for future reference.
- Take new issue description as input as soon as a new incident is reported.
- Retrieves most relevant similar incidents from historical dataset along with previous resolver and resolution details with similarity score.
- Takes user feedback as binary input if the retrieved result was relevant and helpful or not.
- Stores the feedback along with retrieved records to use it further to improve the model.

5 Current Scope

Organization need not depend on local admins round the clock to understand, troubleshoot, and resolve IT issues due to language barriers. For critical applications, where SLA miss can lead to high penalty, the organization can have experts working and resolving issues together from various regions rather than limiting the scope to specific region due to linguistic incompetence.

The scope of work involves.

- Ticket data extraction where support is spread throughout various geographies.
- ITSM ticket data to be considered from single ticket source.
- Data cleansing, data preprocessing, and storing cleansed data.
- Detect ticket language using business service and language field selected by user.
- Use existing IBM Watson translator service to detect language and translate them in English.
- Create a *json* parameter file to store file formats, attribute names, and threshold values to be maintained by users.
- Build information retrieval model using TF-IDF and cosine similarity to extract similar records post machine translation.
- Package the model as microservice to be used as a re-usable asset.
- Take user feedback as binary input and store it with records retrieved for future training and improving retrieved records relevance.
- Response or retrieved records must include ticket priority, creation date, past resolution provided, and resolver details along with similarity score.

If the similarity score is less than the specified threshold, the model should not display those results.

6 Solution Architecture

AIOps framework in ITSM comprises majorly five components as shown in Fig. 2. They are: continuous delivery, continuous operations, environment/infrastructure—cloud, collaboration channel and RPA, and AIOps engine. AIOps engine gets one-way feed from component 1, 2, and 3, and 2-way feed from component 4. In this work, data feed is taken from continuous operations (component 2) and providing the output through collaboration channel (component 4). Feedback is also captured through the same component which further improves the result relevance. The overall flow of work is shown in detail in Fig. 3.

The solution is divided into three major parts, as listed below.

Data extraction and preprocessing:

- Extract the data using ticketing tool REST APIs.
- Fetch business service to identify geography.
- Detect issue language through business service.
- If the detected language is English, then continue with regular preprocessing.
- If the detected language is non-English, then call IBM Watson translator REST API GET method by passing language and text as input parameter.
- Watson translator returns the translated response in English which is stored in a database.

Compute *tf-idf* and Cosine similarity:

- Get new issue description as input and pass-through preprocessing phase.
- Compute *tf-idf* on processed data and the new query.
- Compute cosine similarity on *tf-idf* vectorizer.
- Fetch records with > 0.5 similarity score or based on defined similarity threshold.

Retrieval of relevant records, result evaluation and get user feedback involves:

- Initialize/update weights to each retrieved record such that cumulative weight is 1.
- Retrieve similar issues with historical resolution sorted by weights and score.
- Take user feedback.
- Get the rank at which relevant document appeared.
- Compute mean reciprocal rank and evaluate result.
- Modify weights according to the relevance.

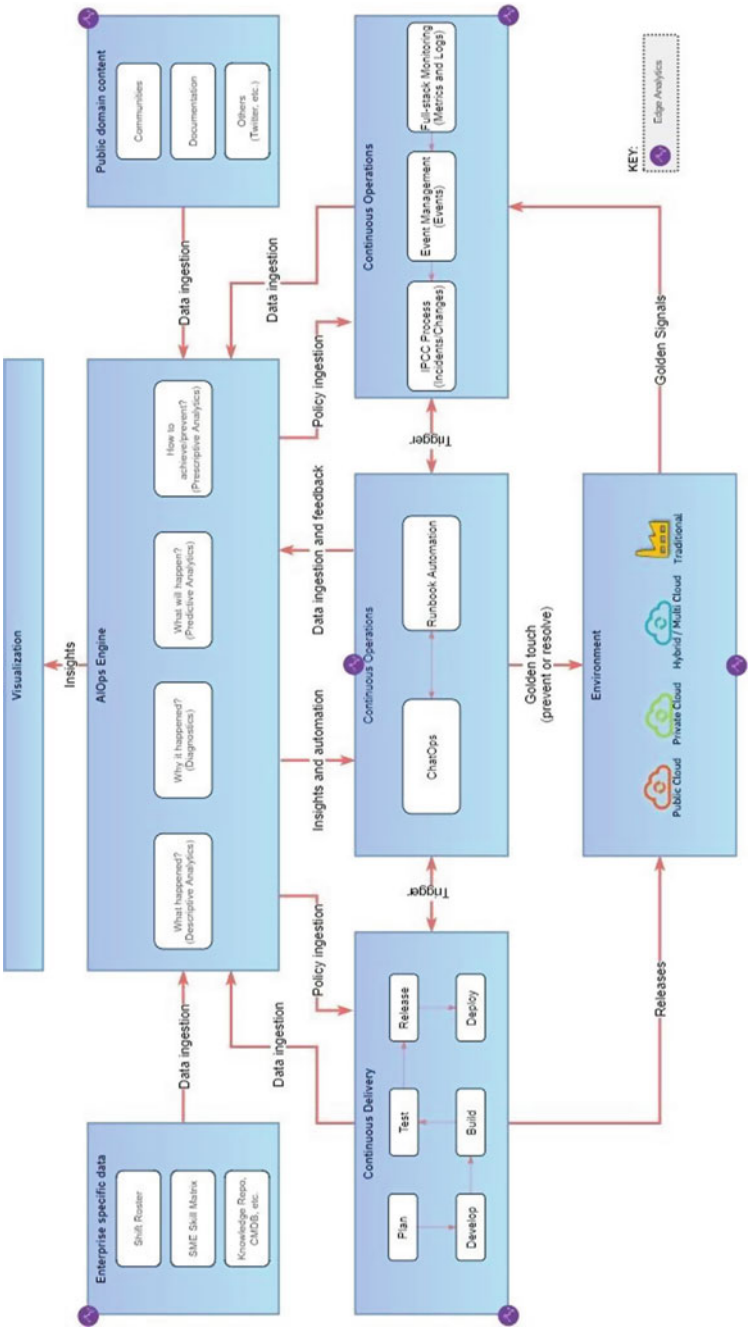


Fig. 2 Generic AIOPs framework

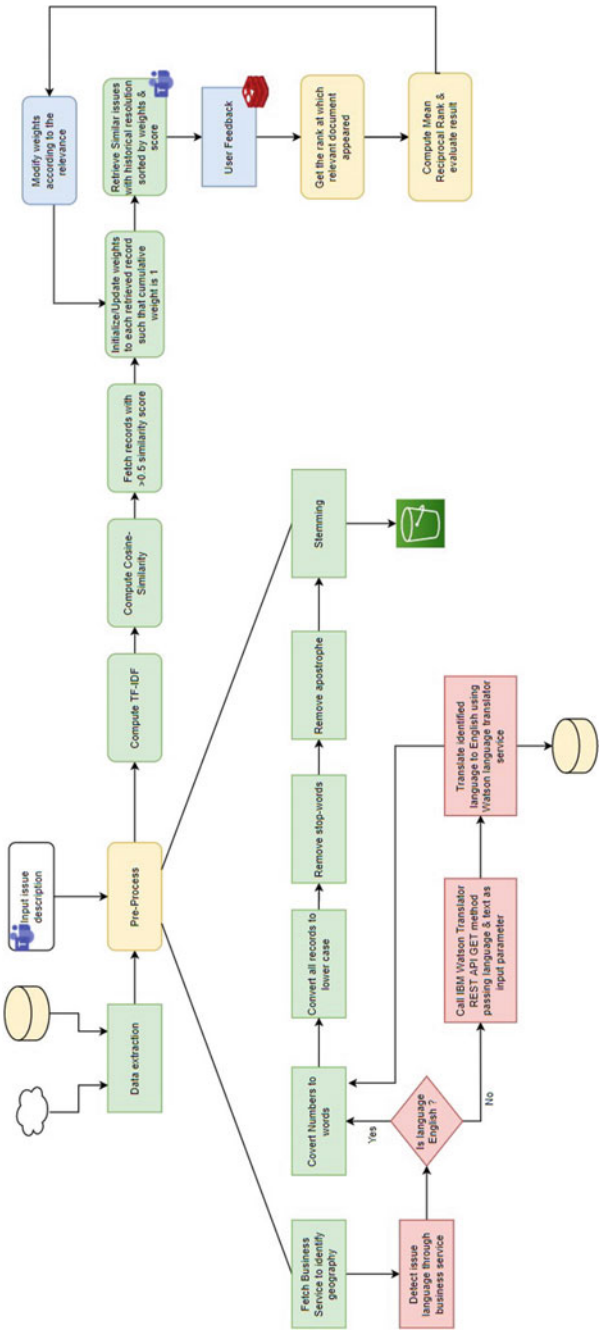


Fig. 3 Data extraction, preprocessing, and computation phases

7 Obtained Results

The tools used for this proposed work are: collaboration platform and slack teams which act as front end to see retrieved results and pass issue description as input parameter. Bot framework is used to call microservice API. Persistent storage S3 is used to store processed data and pickle files. Relational database and MySQL are used to store translated records and user feedback. In-memory cache used to store user feedback temporarily to handle multiple feedbacks from multiple users at the same time. Work is deployed at AWS and Teams. Python, *Nodejs*, and Flask are the software resources. Any Windows/Mac Machine or ECR is hardware used. Further, SRE Team is used to test and evaluate results and provide their valuable feedback as shown in Fig. 4. Application developer is used to integrate microservice with BOT framework. DevOps engineer is to deploy the microservice on cloud, and Watson team is used to expose IBM Watson translator API.

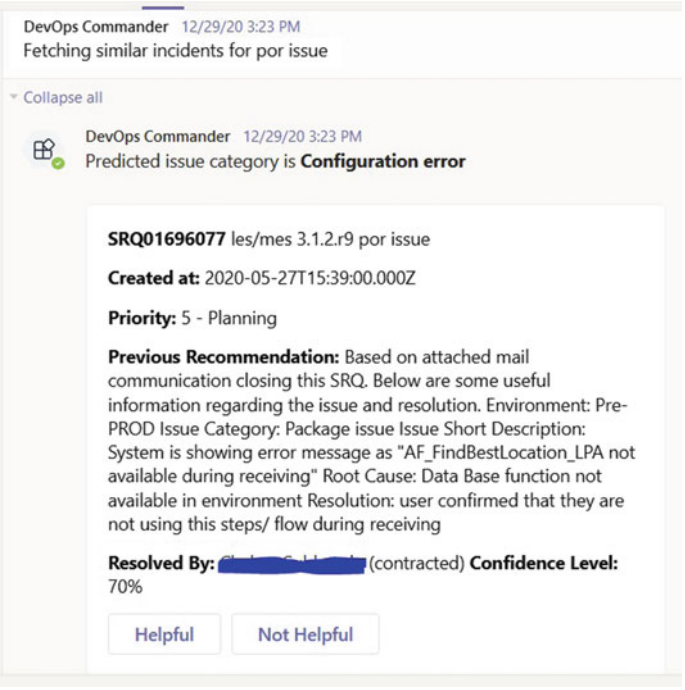


Fig. 4 Microservice API call made through TEAMS

8 Limitations of the Work

The scheduler task should run at the specific time to process new data and generate pickle file based on new dataset on daily basis. The records retrieved sometimes might be totally different than what the user expects and semantically may not be relevant at all. The model which suits best for one customer might be totally irrelevant for another one. It will be realm specific. Maintaining a customer specific is a costly process. The response time to retrieve the similar records will increase if the underlying data size increases. Increase might not be just on number of records but also on text length used.

9 Conclusion and Future Scope

To answer the growing demand for a platform that aggregates a variety of IT data including tickets, logs and alerts, we incorporate similar incident detection of multilingual records in existing AIOps framework. This work will help the SRE to identify re-occurring issues irrespective of language, accelerates root cause investigation using historical resolution details, and permanently fix the problem hampering the application performance. This offers holistic, real-time insights, and an interactive and consultative engagement experience through feedback mechanism.

With this solution, SREs will have an information retrieval system which could retrieve similar historical issues along with its resolution information as soon as a new incident is reported using natural language text so that mean time to resolve an issue reduces significantly and more time is spent on fixing the issue rather than diagnosing it. The user feedback data along with the retrieved record is stored in the database. A scheduler program must be written to send the data from the feedback system to main model to improve the relevance score in a more programmatic manner. A better model could be proposed to handle semantic similarity of issues along with similarity on resolutions details and work notes in the ticket. Advance resolution recommender system can be built through correlation, causality, and pattern identification based on the relevant records retrieved and most likely resolution taken.

References

1. Lee M, Wella ITIL (2011) The maturity of IT service operation in universitas. Multimedia Nusantara, Indonesia. Int J New Media Technol 5:90–94. https://doi.org/10.31937/ijnmt.v5i2.914
2. Obwegeser N, Arenfeldt K, Dam AC, Fenger KH, Silkjaer JV (2020) Aligning drivers, contractual governance, and relationship management of IT-outsourcing initiatives. J Inf Technol Case Appl Res 22(1):40–66

3. Heikkinen S, Jäntti M, Saranto K (2020) Applying continual service improvement practices to study quality of healthcare information system services: a case study. Q Inf Commun Technol 155–168
4. Himi A, Bahsani S, Semma E (2011) The IT Service Management according to the ITIL framework applied to the enterprise value chain. Int J Comput Sci 8
5. Sukmana HT, Wardhani LK, Khairunnisa S, Lee KO, Wati R (2019) ITSM software ranking for small medium enterprises based on ITIL V3 quick win criteria using fuzzy SIR method. Adv Sci Technol Eng Syst 4(2):288–298
6. Wijaya AF, Sitokdana MNN, Hapsari P (2019) Analysis of information technology services management using the ITIL V3 domain service operation framework on SIMDA (Case Study: Boyolali Regency Inspectorate). 100(41):126–131
7. Lubis M, Kusumasari TF, Hakim L (2018) The Indonesia public information disclosure act (UU-KIP): its challenges and responses. Int J Electric Comput Eng 8(1):94–103

Design of Efficient Energy Management Solution for the Internet of Things-Based Smart Microgrid

Mayra Alejandra Pacheco Cunduri, Fabricio Javier Santacruz-Sulca, Diego Ramiro Ñacato Estrella, Danny Velasco Silva, O. Vignesh, and J. N. Swaminathan

Abstract Energy Management Solution (EMS) sounds familiar in recent days for its advantages and solution to monitor and analyse energy consumption. EMS plays a vital role in the energy sector and in the field with bottom-line business priority. This paper proposes an EMS solution to the smart microgrid through an Internet of Things (IoT)-based unified framework. The IoT-based unified framework provides energy efficient optimization and scheduling to the smart microgrid. This work also discusses the energy harvest and energy trading in the rural area where the smart microgrid has been installed. The advantages of the proposed IoT framework are it provides adequate control on the renewable energy resources and optimize the load scheduling based on the energy harvest. To prove the EMS using the developed framework, a simulation analysis has been carried out. The obtained simulation results show the essential requirement of energy management solutions and the optimal performance of the proposed IoT framework.

Keywords Energy management solution (EMS) · Internet of Things (IoT) · Smart microgrid · Energy optimization · Rural power supply

1 Introduction

Taking environment as the major concern, serval energy sectors starting their energy generation through renewable energy resources. When these renewable energy resources are combined to satisfy energy requirements in a community it is known as

M. A. P. Cunduri (✉) · F. J. Santacruz-Sulca · D. R. Ñ. Estrella
Escuela Superior Politécnica de Chimborazo (ESPOCH), Riobamba, Ecuador
e-mail: mayra.pacheco@espoch.edu.ec

D. V. Silva
Universidad Nacional de Chimborazo (UNACH), Riobamba, Ecuador

O. Vignesh
Easwari Engineering College, Chennai, India

J. N. Swaminathan
QIS College of Engineering & Technology, Ongole, Andhra Pradesh 523272, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_9

"microgrid" [1]. Later, when the microgrid generates more power than the requirement they are fed to the grid network with a smart meting service and the microgrid evolved to a smart microgrid [2]. The evolvement brought several benefits to the consumers and the users, making energy trading and load consumption. When taking the smart grid's negative side, the really open issues are viewed as design, procedure and maintenance are the vital problems associated with smart metering. Further, energy management and control when the uncertainties happen can affect the renewable energy resources to create energy unbalance and ruin the devices connected to the smart grid [3, 4]. Several researchers have tried to sort out the energy management problem that exists in the smart grid [5–11]. In [5], discusses the demand response for the commercial user's load by introducing dynamic pricing. They have used a decision support methodology to achieve the same. The decision support methodology is created to figure power demand in the home and empower the client to save energy by suggesting ideal run time plans for machines, given client limitations and season of utilization estimating from the service organization. The timetable is conveyed to the brilliant apparatuses over a self-coordinating home energy organization and executed by the machine control interfaces created in this review. In [6], energy management using multi-agent systems have been developed. The work discusses the application of the multi-agent systems in the power management approach in the grid network is discussed. The study has been analysed through simulation analyses and the result proves the multi-agent systems for energy management. In [7], discusses the information communication technology-based smart grid to improve efficiency and reliability. The work also discusses, the smart grid is relied upon to incredibly get to the next level of the productivity and dependability of future power frameworks with the interest for sustainable power assets. Nonetheless, on the grounds that significant power offices are interconnected through correspondence organizations, the smart grid's network safety is turning into a significant issue where information communication technology provides a suitable solution. In [8], discusses the solution to reduce the peak to average ratio and the consumer delay time by a management technique applied to the smart appliances. In [9], discusses the wireless sensor networks for various energy harvesting technologies to optimize energy saving applications. The research also explores the issues on energy effectiveness for sensor networks utilizing energy gathering innovation are then talked about. Finally, the work also provides an ideal energy-based strategy for a sun-based fueled sensor hub that utilizes a rest and wakeup system for energy protection. In [10], develops a framework of smart grid energy management solution that includes different functions and could manage and provide next-generation monitoring, investigation, and regulatory functions. The developed approach could provide an efficient solution to the smart grid for business energy management in the critical part of the future smart grid. In [11], discusses the rapid urbanization that took place in the recent past and the need for a sustainable solution for energy, transport, environment, and life quality. The work developed the Internet of Things (IoT) as a solution for smart utilization and requirement. Hence, IoT could play a major role in smart appliances and could make changes in the digital world by its advantages on

web and device interface. In [12], discusses the IoT-based smart metering infrastructure in the power grid to enhance the power quality and reliability of the grid network. In the developed approach, the existing remote correspondence advancements and applicable information directing calculations are entirely talked about in the IoT. In [13] discusses, the IoT for computation analysis and communication in the smart grid network. The smart grid with IoT could provide additional communication and control to the smart grid. In [14], explore the correspondence prerequisites of the savvy network and present all IoT conventions and their determinations.

2 System Description

The studied energy management solution has been applied on the smart microgrid shown in Fig. 1. The smart microgrid system comprises of two renewable energy resources Permanent Magnet Generator (PMG)-based wind energy conversion system and Photo Voltaic (PV)-based solar energy system. Where the wind energy conversion system delivers a maximum power of 1 kW and PV system generates a max power of 1 kW. The wind energy conversion system and PV system has been modelled using the Eqs. (1) and (2) given by,

$$P_m = \frac{1}{2}\rho C_p(\lambda, \beta)\mathrm{AW}^3 \tag{1}$$

$$I_{\mathrm{pv}}(V_{\mathrm{pv}}) = I_{\mathrm{SC}}\left[1 - C_1\left\{e\left(\frac{V_{\mathrm{pv}}}{C_2 U_{\mathrm{oc}}}\right) - 1\right\}\right] \tag{2}$$

The wind energy conversion system is connected to a 24 V DC bus system through an AC/DC converter. Meanwhile, PV system is connected to the 24 V DC bus system

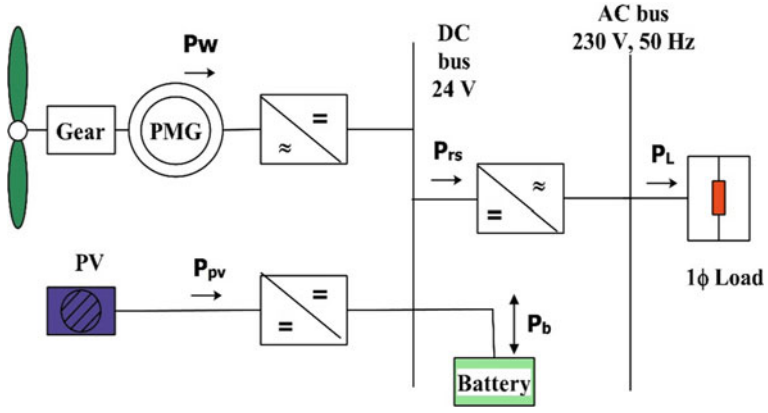


Fig. 1 Single line diagram of the studied smart microgrid

through an DC/DC optimal power tracking boost converter. This converter helps to achieve the optimal power and regulates the DC power fed to the DC bus system. A lead acid battery of 600 Ah, 24 V is connected to the DC bus through a charge controller and the battery is modelled using the Eq. (3).

$$I_b = \frac{\left[V_{\mathrm{OC}} - \sqrt{V_{\mathrm{OC}}^2 - 4(R_a + R_b)p_b}\right]}{2(R_a + R_b)} \tag{3}$$

Further, the power generated through the renewable energy sources and battery fed to the DC/AC converter is given by,

$$P_{\mathrm{rs}} = P_w + P_{\mathrm{pv}} + P_b \tag{4}$$

The power Prs is converted and fed into the AC bus of 230 V, 50 Hz. Since, the generated power is minimum only single-phase loads are considered in the developed smart grid. A non-linear load has been connected and analysed in the studied system.

3 Energy Management Domain Structure

A dreadful requirement for Energy Management Solution (EMS) has arisen all around the globe due to the raise of expansion of energy requirement and its utilization. The EMS domain structure is shown in the Fig. 2. The EMS structure encapsules two solution model such as, (i) Energy efficient solutions and (ii) Energy harvesting solutions. These solution models derive the optimal solution to the EMS. Further, the energy efficient solutions provide possible requirement like cloud approach, scheduled optimization and energy consumption. Meanwhile, the energy harvesting solution possibly provides scheduled generation, energy routing and trade-off. The EMS structure developed will act as a decentralized management system that can control the generation, optimum requirement, load scheduling, energy routing and trade-off from anywhere using these options.

4 Proposed IoT-Based Smart Wearable Tags

The function of the developed IoT-based EMS system is shown in Fig. 3. The developed IoT EMS performs four functions namely, economic operation, energy efficiency, load scheduling and interface, and source selection and scheduling. The main objective of the developed IoT-based EMS is energy consumption and economic operation. The concept also brings demand side management with energy trading. The energy trading is an idea of trading the produced electrical energy from the own environmentally friendly power assets to the miniature lattice associated houses,

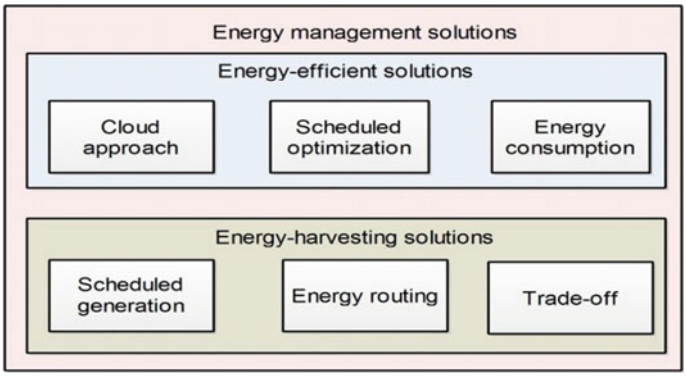


Fig. 2 Energy management solution structure

workplaces and plants. The energy trading is conceivable between a local area or another local area close by. The energy trading permits the clients to choose to whom the power can be sold, and from whom they can buy power. The data connected with the trading will be put away in the changeless record that can be just gotten to by the designer and one with certifications. The upside of created framework is that it is critical, during the low interest condition the energy produced through sustainable power assets will draw more energy and when acquainted with the primary network that can influence the lattice, and cause interference. The framework additionally give security to the smart gird environment. The developed IoT EMS uses linear and integer programming analysis type to find suitable solution to the objective.

5 Results and Discussion

To validate the proposed IoT-based EMS for IoT operated smart microgrid an extensive simulation analysis has been carried out using the Machine Learning Studio (classic). The system considered for the study is I3 processor with 6 GB RAM, Windows 10 OS. Initially the experiment setup has been modelled using the mathematical expression and single line diagram discussed in the section two. Further, voltage, current and power sensor has been configured to an Arduino uno with ESP8266 module. It is to be noted the internet service must be continually provided to read the data and write the data in the IoT channel. The data fetched from the module is transferred to the cloud server using the ESP8266 module. Those data which has been recorded in the IoT server helps to analyse the energy production and demand. The power generated by renewable energy sources in the IoT-based smart microgrid is shown in Fig. 4. From the figure it can be seen that the power generated by PV, wind and total power is depicted in W for a period of 3 days. It can be overserved that PV generated an average power of 60–80 W per day due to low sun shine and wind generates an average power of 400 W and the total power

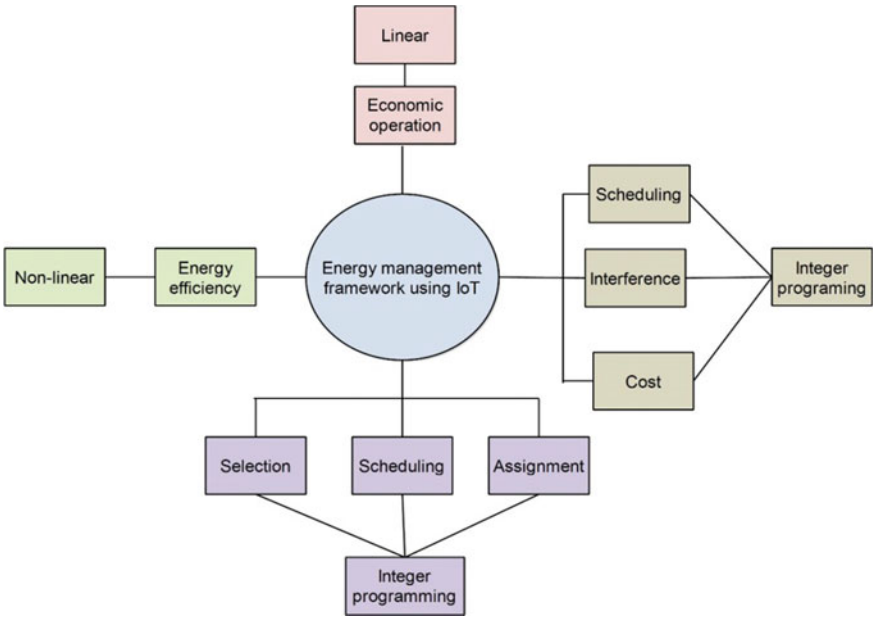


Fig. 3 Function of IoT-based EMS

generated by the renewable energy sources is about 500 W based on sun radiation and wind flow. The validation of the developed EMS working with IoT is carried out based on two cases. In first case EMS on electricity tariff and second case on energy optimization. Figure 5 depicts the load pattern observed for electricity tariff. It can be observed that the tariff is uneven based on the saturation in the energy production and load demand. The tariff in the EMS is fixed in such a way that when the energy production is higher than the demand the tariff will be lower and in case when the energy demand is higher than the production the tariff will be high. In such cases, the IoT smart grid communicates with the grid network to satisfy the load demand. The energy transaction carried out using the IoT-based EMS is shown in Fig. 6. From the results it can be observes that traditional method has high impact on transaction.

The EMS can provide an optimal power transaction and energy trading. The EMS monitors the energy trading between the consumer-consumer or power sector-consumer. The EMS bids the tariff according to the load generation and demand. The consumer will be provided a choice to choose his one supplier and bid the supplier based on the requirement and tariff. Thus, the EMS allows the user to choose their own supplier at low cost. In the second case of energy optimization analysis, Fig. 7 depicts the EMS energy optimization obtained for the IoT-based smart microgrid. From the figure it can be observed that the green cluster shows the grid power, blue cluster shows the load demand, and orange cluster shows the energy generated in the studied smart microgrid. In case, when the load demand is higher than the energy generated, the EMS optimize the load demand and gets satisfied from the grid network. Further

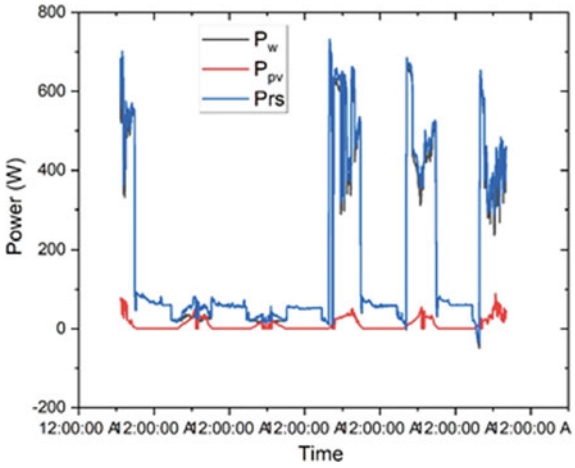


Fig. 4 Power generated in the IoT-based smart microgrid

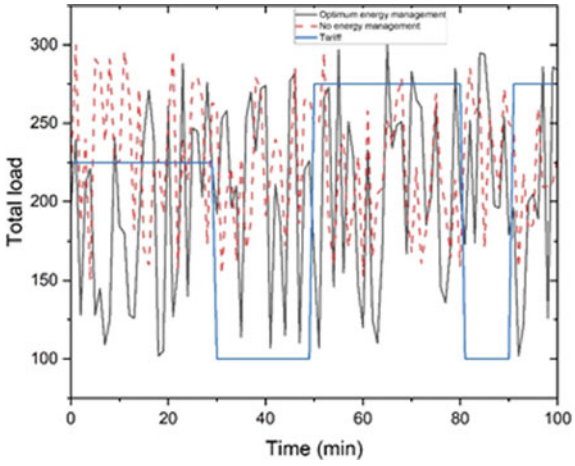


Fig. 5 Load pattern for electricity tariff

the case, when the load demand is lesser than the energy generated. The EMS system optimistically satisfy the load and excess power is energy traded. Thus, EMS provide suitable optimization to the smart microgrid.

6 Conclusion

Energy management solution for the Internet of Things (IOT)-based smart microgrid has been studied in this paper. In the analysis, the energy management solution is classified into two distinct levels such as energy efficient solutions and energy harvesting solutions. The analysis has been carried out in various directions to assess

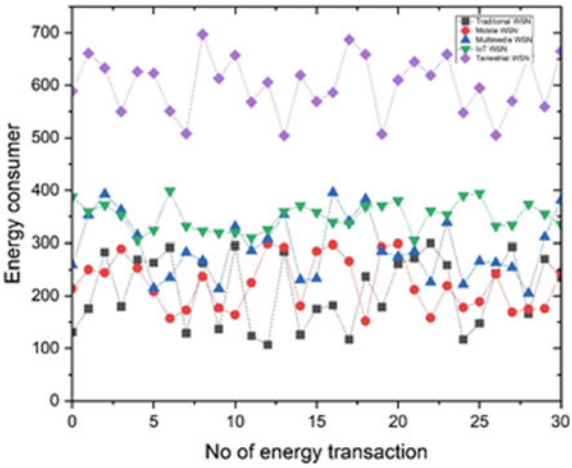


Fig. 6 Energy transaction using the EMS

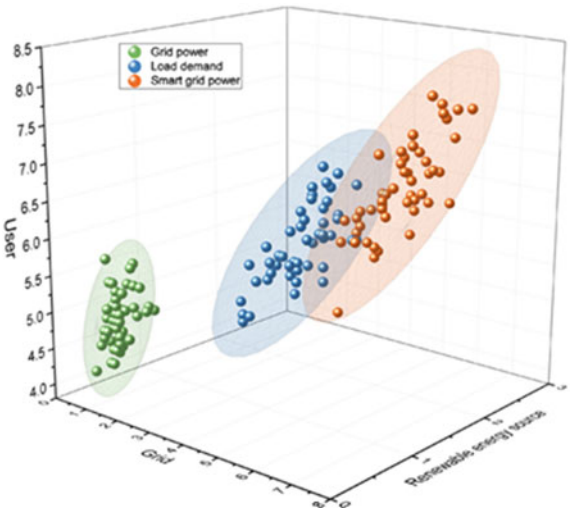


Fig. 7 Optimization using the EMS

the energy efficient solution and energy harvesting solution using the EMS for the IoT-based smart microgrid. The study conducted identifies the optimal and efficient scheduling in the smart microgrid to reduce the cost of operation and electricity bill. Further, the simulation results have been presented to support the working of EMS in the IoT-based smart microgrid.

References

1. Sitharthan R, Geethanjali M, Pandy TKS (2016) Adaptive protection scheme for smart microgrid with electronically coupled distributed generations. Alex Eng J 55(3):2539–2550
2. Soundarya G, Sitharthan R, Sundarabalan CK, Balasundar C, Karthikaikannan D, Sharma J (2021) Design and modeling of hybrid DC/AC microgrid with manifold renewable energy sources. IEEE Canadian J Elec Comput Eng 44(2):130–135
3. Judge MA, Khan A, Manzoor A, Khattak HA (2022) Overview of smart grid implementation: frameworks, impact, performance and challenges. J Energy Storage 49:104056
4. Ramachandran S (2020) Applying AI in power electronics for renewable energy systems [expert view]. IEEE Power Electron Mag 7(3):66–67
5. Ozturk Y, Senthilkumar D, Kumar S, Lee G (2013) An intelligent home energy management system to improve demand response. IEEE Trans Smart Grid 4(2):694–701
6. Lagorse J, Paire D, Miraoui A (2010) A multi-agent system for energy management of distributed power sources. Renew Energy 35(1):174–182
7. Kim SM, Lee T, Kim S, Park LW, Park S (2019) Security issues on smart grid and blockchain-based secure smart energy management system. In: MATEC Web of conferences, vol 260. EDP Sciences, p 01001
8. Shuja SM, Javaid N, Rafique MZ, Qasim U, Khan RFM, Butt AA, Hanif M (2019) Towards efficient scheduling of smart appliances for energy management by candidate solution updation algorithm in smart grid. In: International conference on advanced information networking and applications. Springer, Cham, pp 67–81
9. Niyato D, Hossain E, Rashid MM, Bhargava VK (2007) Wireless sensor networks with energy harvesting technologies: a game-theoretic approach to optimal energy management. IEEE Wirel Commun 14(4):90–96
10. Park K, Kim Y, Kim S, Kim K, Lee W, Park H (2011) Building energy management system based on smart grid. In: 2011 IEEE 33rd international telecommunications energy conference (INTELEC). IEEE, pp 1–4
11. Ejaz W, Naeem M, Shahid A, Anpalagan A, Jo M (2017) Efficient energy management for the internet of things in smart cities. IEEE Commun Mag 55(1):84–91
12. Al-Turjman F, Abujubbeh M (2019) IoT-enabled smart grid via SM: an overview. Futur Gener Comput Syst 96:579–590
13. Bekara C (2014) Security issues and challenges for the IoT-based smart grid. Procedia Comput Sci 34:532–537
14. Tightiz L, Yang H (2020) A comprehensive review on IoT protocols' features in smart grid communication. Energies 13(11):2762

Building Secured Software Defined Networks by Analyzing Anomaly Detection Algorithms on Various Attacks

R. Presilla, Jagadish S. Kallimani, and Rajesh Eswarawaka

Abstract This work is to highlight the various methods and approaches to conduct anomaly detection in application log messages in a typical software defined networks. Work also compares the precision, accuracy and sensitivity of various supervised and unsupervised methods of anomaly detection. Typically, in large scale software defined networks, anomaly detection is an important entity in managing it. Log records are widely used for anomaly detection to record system runtime information. As part of this work, guidelines for adopting this work and scope for future works are proposed. The review can assist in anomaly detection method and avoid doing any redundant works. The steps involved in the process are: log collection, log parsing, feature extraction, and anomaly detection. More specifically, reviewed anomaly detection methods include supervised methods (Logistic Regression, Decision Tree, and Support Vector Machines) and unsupervised methods (Log Clustering, Principle Component Analysis, and Invariant Mining). They are being discussed in detail along with comparison results.

Keywords Anomaly detection · Application logs · Software defined networks · Machine learning · Neural networks · Supervised methods · Unsupervised methods · Malicious attacks · Log data

R. Presilla
Department of Computer Science and Engineering, M S Ramaiah Institute of Technology, Bangalore and Affiliated to Visvesvaraya Technological University, Belagavi, India

Department of Artificial Intelligence and Machine Learning, M S Ramaiah Institute of Technology, Bangalore, Karnataka, India

J. S. Kallimani (✉) · R. Eswarawaka
Department of Artificial Intelligence and Machine Learning, AMC Engineering College, Bangalore, India
e-mail: jagadish.k@msrit.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_10

1 Introduction

DevOps and Artificial Intelligence (AI) are interdependent on each other in many aspects. DevOps is a business-driven approach and AI is the technology for enhanced functionality. Under the influence of AI, DevOps teams can test, code, release, and monitor software systems efficiently [1, 2]. AI can also improve automation, resolve issues, and improve collaboration between teams [3]. Given below are few ways AI is transforming DevOps:

- Automated software testing.
- Defect identification.
- Finding potential issues before we kick the testing.
- Automated management of issues found in the testing process in real time.
- Automated resolution of issues discovered in build, deployment, and testing.
- Automated documentation and easy auditing of issues.
- Effective managing of QC and security testing bringing better development awareness, collaboration, and effectiveness.
- Managing consistent configurations across all environments.
- Simulated user-level test cases driven by AI.

Process optimization assists in fine tuning of the change management processes and procedures for moving applications from devices to product environments leveraging historical data [4]. It acts as smart management of code freeze and un-freeze. It assists in optimizing the development cycle and providing proper planning based on data from delivery history of previous projects. Alert Management prioritizes response/time and assigning alerts to proper teams based on factors such as past behavior, source of the alerts, and volume. Automated compliance checks across applications/integrations, vendors, and Environments [5].

Automation of routine tasks within the workflows involves: capturing information on build failure/success and generating and assigning automated alerts to all stakeholders, and orchestrating complex pipelines. It assists in automated classification and root cause analysis of build/deployment data [6]. It uses predictive analytics to flag potential areas of concern for future builds. Assignment of bug fixes and issues to the proper people, based on build/commit/deploy data (even from previous projects) are involved. Embedding security testing within the DevOps cycle and optimizing and securing Infrastructure as code concepts are also additional responsibilities which involves repeatability and security [7].

Improved collaboration tasks are: facilitate continues feedback loops by proactively identifying potential and real issues early on and making recommendations to address timely, facilitate more effective collaboration between the devices and teams. Automatic posting of issues and tagging proper people in various collaboration tools (Slack) and data correlation across platform and tools by analyzing data streams from various heterogeneous systems and tools to find correlations and create a holistic view of new deployments, production issues, or application's health are

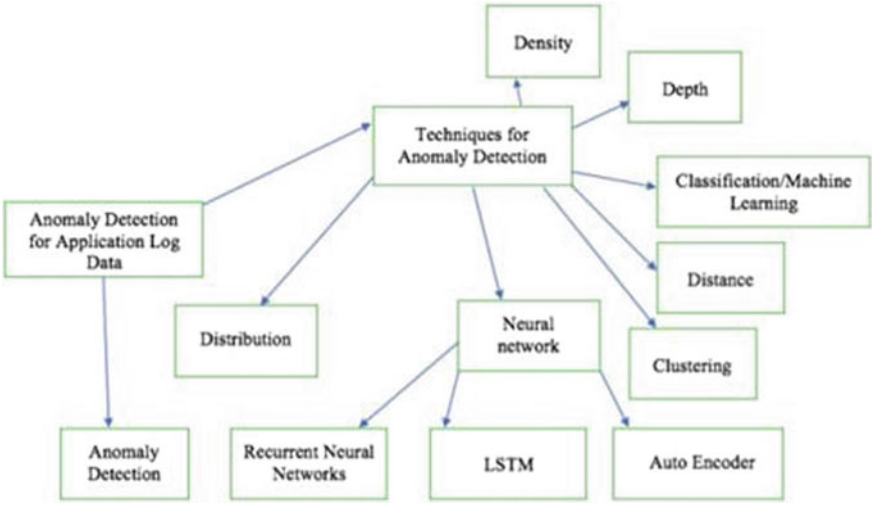


Fig. 1 Techniques for anomaly detection

also involved. Personality awareness for team building and interactions are further taken care of.

Applications generate huge amounts of log data that are time stamped. They represent every single system and user event for all users of the application [1]. Continuous generation of data throughout the lifetime of an application is huge, time consuming and high complex in nature [8]. Figure 1 shows the typical setup of anomaly detection technique.

2 Experimental Setup

The above framework primarily involves—collection of logs, parsing of lags, extraction of features, and detection of anomaly as per Fig. 2. Various states of the system and generated runtime information are timely recorded by logs. Temporal data and description of the events are recorded in logs. Several applications including anomaly detection are being processed through these log data. For instance, the distributed file system logs on a commercial platform are shown for better understanding.

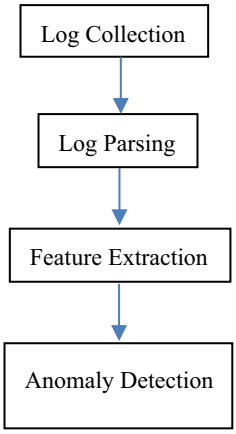


Fig. 2 Anomaly detection framework

LineId	Date	Time	Pid	Level	Componer	Content	EventId	EventTemplate
1	81109	203518	143	INFO	dfs.DataN	Receiving l	E5	Receiving block <*> src: /<*> dest: /<*>
2	81109	203518	35	INFO	dfs.FSNam	BLOCK* N	E22	BLOCK* NameSystem.allocateBlock:<*>
3	81109	203519	143	INFO	dfs.DataN	Receiving l	E5	Receiving block <*> src: /<*> dest: /<*>
4	81109	203519	145	INFO	dfs.DataN	Receiving l	E5	Receiving block <*> src: /<*> dest: /<*>
5	81109	203519	145	INFO	dfs.DataN	PacketRes	E11	PacketResponder <*> for block <*> terminating
6	81109	203519	145	INFO	dfs.DataN	PacketRes	E11	PacketResponder <*> for block <*> terminating
7	81109	203519	145	INFO	dfs.DataN	Received b	E9	Received block <*> of size <*> from /<*>
8	81109	203519	145	INFO	dfs.DataN	Received b	E9	Received block <*> of size <*> from /<*>
9	81109	203519	147	INFO	dfs.DataN	PacketRes	E11	PacketResponder <*> for block <*> terminating
10	81109	203519	147	INFO	dfs.DataN	Received b	E9	Received block <*> of size <*> from /<*>
11	81109	203519	29	INFO	dfs.FSNam	BLOCK* N	E26	BLOCK* NameSystem.addStoredBlock: blockMap updated: <*> is added to <*> size <*>
12	81109	203519	30	INFO	dfs.FSNam	BLOCK* N	E26	BLOCK* NameSystem.addStoredBlock: blockMap updated: <*> is added to <*> size <*>
13	81109	203519	31	INFO	dfs.FSNam	BLOCK* N	E26	BLOCK* NameSystem.addStoredBlock: blockMap updated: <*> is added to <*> size <*>

def _split_data(x_data, y_data = None, train_ratio = 0, split_type = 'uniform')
def load_HDFS(log_file, label_file = None, window = 'session', train_ratio = 0.5, split_type = 'sequential', save_csv = False, window_size = 0):
"""" Load HDFS structured log into train and test data
Arguments
log_file: str, the file path of structured log.
label_file: str, the file path of anomaly labels, None for unlabeled data
window: str, the window options including 'session' (default). train_ratio: float, the ratio of training data for train/test split
split_type: 'uniform' or 'sequential', which determines how to split dataset. 'uniform' means
to split positive samples and negative samples equally when setting label_file 'sequential'
means to split the data sequentially without label_file. That is, the first part is for training, while the second part is for testing.
Returns
(x_train, y_train): the training data (x_test, y_test): the testing data
class FeatureExtractor(object)

def fit_transform(self, X_seq, term_weighting = None, normalization = None, oov = False, min_count = 1):
Fit and transform the data matrix
Arguments
X_seq: ndarray, log sequences matrix term_weighting: None or 'tf-idf' normalization: None or 'zero-mean' oov: bool, whether to use OOV event
min_count: int, the minimal occurrence of events (default 0), only valid when oov = True
Returns
X_new: The transformed data matrix
def transform(self, X_seq):
Transform the data matrix with trained parameters
Arguments
X: log sequences matrix term_weighting: None or 'tf-idf'
Returns
X_new: The transformed data matrix """
def init (self, n_components = 0.95, threshold = None, c_alpha = 3.2905): """
The PCA model for anomaly detection
def fit(self, X):
def predict(self, X):
def evaluate(self, X, y_true)

BlockId	Label
blk_-1608999687919862906	Normal
blk_7503483334202473044	Normal
blk_-3544583377289625738	Anomaly
blk_-9073992586687739851	Normal
blk_7854771516489510256	Normal
blk_1717858812220360316	Normal
blk_-2519617320378473615	Normal
blk_7063315473424667801	Normal
blk_8586544123689943463	Normal
blk_2765344736980045501	Normal
blk_-2900490557492272760	Normal

3 Experiments and Analysis

As per the proposed work, the experiment is conducted and results are recorded. Various analysis on the obtained results are conducted and they are presented in this section. Typically, we consider the window size of one hour and step size of six hours.

Like supervised methods, experiments on different settings of window size and step size to explore their effects on accuracy. It is noticed and recorded as per the above Table 1. This has been experienced as per the workflow followed in Fig. 3.

Table 1 Obtained results

	HDFS					
	Train			Test		
Model	Precision	Recall	F1	Precision	Recall	F1
LR	1.000	0.365	0.535	0.986	0.433	0.602
Decision tree	1.000	0.365	0.535	0.985	0.427	0.596
SVM	1.000	0.365	0.535	0.986	0.433	0.602
PCA	0.967	0.186	0.312	0.966	0.363	0.528
Invariants mining	0.947	0.305	0.462	0.973	0.559	0.710
Clustering	1.000	0.359	0.528	0.967	0.561	0.710

	HDFS		
Model	Precision	Recall	F1
LR	0.955	0.911	0.933
Decision tree	0.998	0.998	0.998
SVM	0.959	0.970	0.965
PCA	0.975	0.635	0.769
Invariants mining	0.888	0.945	0.915
Clustering	1.000	0.720	0.837

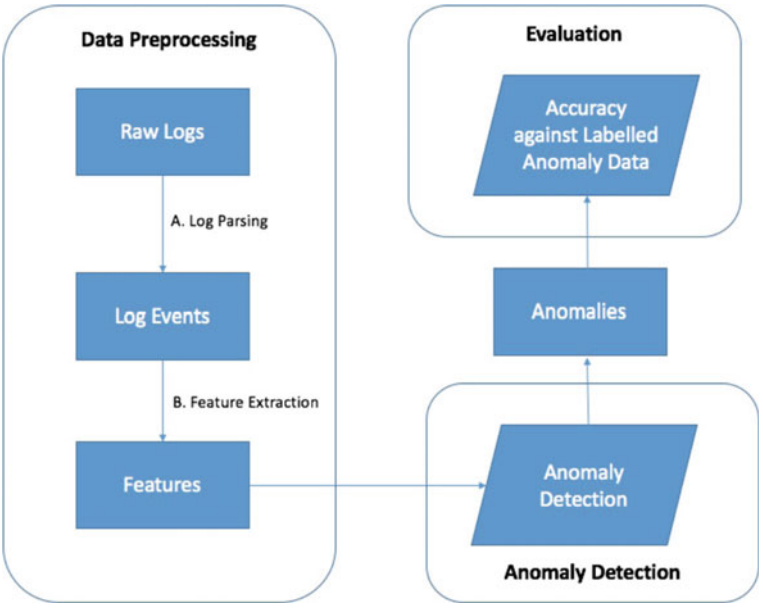


Fig. 3 Anomaly detection workflow

4 Efficiency of the Proposed Model

By considering various log files of different sizes, efficiency of anomaly detection methods is evaluated. Anomalies are detected with less time in supervised methods and more time in unsupervised methods. As the log files size increases, anomaly detection also consumes huge time except log clustering. Other methods require advanced optimization methods to speed up the process.

5 Conclusions and Future Works

Log analysis and anomaly detection techniques are used and analyzed in large scale in recent times. An optimal anomaly detection method is yet to evolve. In many situations, researchers re-design customized anomaly detection methods. This is caused because of short comes in exhaustive reviews and knowledge about optimal methods. Machine learning models have given wide scope of developing novel methods for anomaly detection as per Fig. 4. Usually researchers fail to identify the anomalies in some real time situations. New techniques and approaches needs to be evolved to identify the correct behavior of anomalies. Presently logs are generated in large volumes by the available systems. Hence, it is challenging to deal with huge log files currently. In the context of big data analytics, generation of analytic tools and detection techniques are the need of the hour.

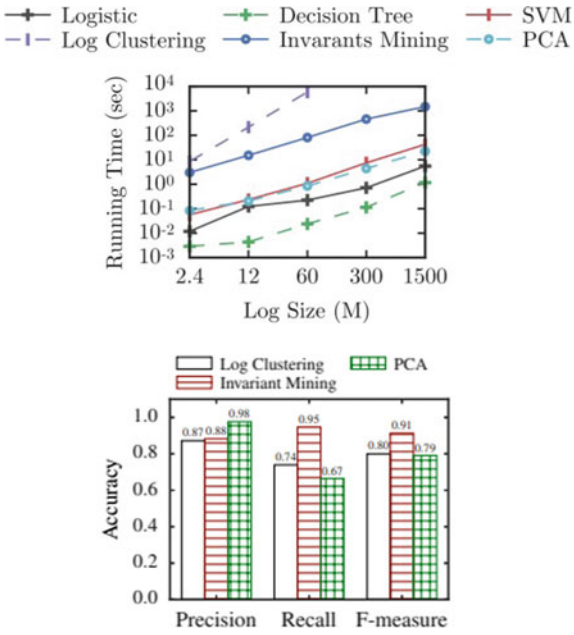


Fig. 4 Obtained results as per the considered metrics

High recall rate is always difficult during anomaly detection. It is highly difficult to identify the rare and heterogeneous anomalies. In many real time situations, legal instances are identified as anomalies and real anomalies are not identified. Multi-dimensional and dependent data poses huge challenge in identifying the anomaly. Hence, development of methods to detect labeled and unlabeled anomalies are expensive tasks and yet to be explored. Concept of conditional and group anomalies increases the complexity in detecting anomalies. These challenges are to be explored by researchers in future.

References

1. He S, Zhu J, He P, Lyu MR (2016) Experience report: system log analysis for anomaly detection. In: 2016 IEEE 27th international symposium on software reliability engineering (ISSRE), pp 207–218. https://doi.org/10.1109/ISSRE.2016.21
2. He P, Zhu J, He S, Li J, Lyu R (2016) An evaluation study on log parsing and its use in log mining. In: DSN'16: proceeding of the 46th annual IEEE/IFIP international conference on dependable systems and networks
3. Lanzaro R, Natella S, Winter D, Cotroneo, NS (2014) An empirical study of injected versus actual interface errors. In: ISSTA'14: proceeding of the 2014 international symposium on software testing and analysis. ACM, pp 397–408
4. Liang Y, Zhang Y, Xiong H, Sahoo R (2007) Failure prediction in IBM bluegene event log. In: ICDM'07: proceeding of the 7th international conference on data mining
5. Lin Q, Zhang H, Lou JG, Zhang Y, Chen X (2016) Log clustering based problem identification for online service systems. In: ICSE'16: proceeding of the 38th international conference on software engineering
6. Lou J, Fu Q, Yang S, Xu Y, Li J (2010) Mining invariants from console logs for system problem detection. In: ATC'10: proceeing of the USENIX annual technical conference
7. Makanju, Zincir-Heywood A, Milios E (2009) Clustering event logs using iterative partitioning. In: KDD'09: proceeding of international conference on knowledge discovery and data mining
8. Zhao Z, Xu C, Li BA (2021) LSTM-based anomaly detection model for log analysis. J Sign Process Syst 93:745–751. https://doi.org/10.1007/s11265-021-01644-4

Ultra-Sensitive Optical Sensor to Detect Single Waterborne Bacterium

Afzal Shaikh, Manju Devi, and Shaista Shaikh

Abstract Water safety is a serious problem, particularly in highly populated cities, because pathogenic microorganisms in drinking water can trigger deadly outbreaks. Bacterial contamination in drinking water, like *Shigella flexneri*, *Vibrio cholera*, *Salmonella enterica*, *E. coli*, and others, is a major source of pathogenic pollutants. The current study proposes a photonic crystal (Ph.C.) micro square ring resonator (MSRR) sensor that is small in size and has a fast output response time to detect the signature of single waterborne bacteria. The bacterial contamination in water varies the refractive index and in turn resonance wavelength shift at the throughput port is observed. The transmission spectrum is determined by using FDTD simulation of Ansys Lumerical software for the suggested structure and the impurity is determined by examining the resonance peak wavelength shift. The proposed structure is ultra-sensitive and compact in size having sensitivity of almost 946 nm/RIU and the structure would be useful for a variety of sensing purposes.

Keywords Photonic crystal · Micro ring resonator · FDTD · Sensor · *E. coli* · *Vibrio cholera*

1 Introduction

Scientific measurements are indeed essential to periodically assess the contamination of water from the several supply sources. Pollutants from sewage entering the borehole, through corroded pipelines, bacterial growth on stagnated water, the residual chlorine levels in water are limited or feeble can be the source of contamination. The procedures for water quality monitoring are segregated into physical, chemical,

A. Shaikh (✉)
Anjuman-I-Islam's Kalsekar Technical Campus, Panvel, India
e-mail: afzal.aiktc@gmail.com

M. Devi
The Oxford College of Engineering, Bengaluru, India

S. Shaikh
Anjuman-I-Islam's A. R Kalsekar Polytechnic, Panvel, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_11

bacteriological, and microscopic categories. It is possible to identify the traces of several microorganisms in water and harmful chemicals. The approaches of separation and detailing are mostly very tedious processes and time-taking. It is looking like it will be difficult to measure drinking water contamination for a pollutant on the fly [1]. The present context of work is mainly related to chemical analysis, which detects toxic chemicals and the detection of harmful organisms.

This paper presents an alternate approach to detecting pollutants in water using an optical sensor. The work shows design, modeling, and simulation of optical sensor using photonic crystal. A silicon photonic crystal-based micro square ring resonator structure is analyzed and designed using line and point defects in order to create a sensor that will be used for the detection of single waterborne bacterium like *Shigella flexneri*, *Vibrio cholera*, *Salmonella enterica*, *E. coli*, and others. The Lumerical FDTD simulation tool is used to design and model the proposed structure. The performance parameters sensitivity, quality factor, and fast response are used to evaluate the proposed structure.

2 Literature Survey

Miniaturization is the focus of current research. Using light technology, photonic crystal-based biosensors can now detect various bio-molecules, bacteria, viruses, and contaminants with high accuracy.

Many novel concepts and applications in one, two, and three dimensional, as well as structures such as dielectric, metallic, and acoustic, are being researched today. E. Yablonovitch explained the initial motivations for this study, which sprang from the need for a photonic band gap in quantum optics. Several practical and theoretical research were motivated by this necessity. This paper demonstrated how to dope a photonic semiconductor to create small electromagnetic cavities [2]. Granum and Lund presented about *Bacillus cereus* a form of bacteria and its water poisoning toxins. This is one of the early works of waterborne bacteria identification [3]. Tuminello et al. were the one to investigate the optical characteristics of *Bacillus subtilis* spores ranging in size from 0.2 to 2.5 num [4]. These research efforts led the foundation of using optical properties of bacteria for identification in water.

Sharan et al., worked on contaminants in drinking water using Ph.C. This identifies the impurities such as magnesium, calcium, iron, and cobalt [5]. Further research on salt in water resulted in further work, such as Lavanya et al., who proposed a concept of an optical sensor for detecting brininess in water [6]. Further, enhancement of the work using AdaBoost algorithm for analysis by S. G, Vivek et al. is interesting (S. G, Vivek et al. 2015). Praveen et al. created a Ph.C.-based nanoscaled sensor for detecting typhoid-causing chemicals in water. This work is a path breaking effort to help primary healthcare staffs in early detection of typhoid [7].

Sharma and Kalyani conducted early study on developing four-channel nano cavities linked Ph.C.-based biosensors for detection of bacteria in water. This is one of the unique efforts where multi-channel approach was experimented [8]. Roy and

Sharan's recent work on the creation of ultrahigh sensitivity biosensors to detect *E. coli* in water is in-depth research on waterborne *E. coli* detection. The research focuses on the use of SPR-based biosensors to detect *E. coli* [9].

Inspired by the advantage of using Ph.C. techniques for biosensing and optimization of outcome by machine learning, S. A. Nehal et al. published their work on AI for bacteria detection. This work uses the unique spectrum signature of the waterborne bacteria to identify water contamination [10].

A review of the literature was conducted, which included different studies and research contributions in the field of photonic crystal-based optical sensors. So, the main goal of this paper is to design and simulate a photonic crystal-based optical sensor capable of detecting waterborne bacterium as well as analyzing and comparing parameters such as sensitivity, quality factor, fast response, and so on.

3 Proposed Structure for Waterborne Bacteria Analysis

The Ph.C. structure used for sensor design is a square lattice structure composed of silicon (Si) rods placed in a silica slab (SiO_2) (see Fig. 1). Inside the Ph.C. structure, an MSRR structure is created, forming a line and a point defect. Through numerical modeling, the optimal lattice constant (a) and silicon rods radius (r) are 540 nm and 100 nm, respectively, yielding a r/a ratio of value 0.185. The ring's center is composed of 3×3 rods and arranged in a square pattern to improve responsiveness and sensitivity. The four Si rods are placed at the four corners of the ring to reduce losses such as scattering loss [11]. In this reported work, the proposed structure is placed in the water sample, based on the different bacteria contamination the effective refractive index of the water sample gets change. As a result, the light encounters change in refractive index surrounding the inner ring of the proposed structure, causing the resonance peak wavelength to vary.

The field pattern in a 2D Ph.C. proposed structure is as shown (see Fig. 2). In the TM mode, the field is perpendicular to the plane of periodicity for rods in air arrangement, i.e., parallel to the rods, while in the TE mode, the plane of periodicity runs around the holes.

As a result, for rod-type structures, TM mode is selected over TE mode polarizations. Furthermore, TM mode polarizations produce wider band gaps in rod type structures than TE mode polarizations. The TM mode provide large band gap which is because of the strong RI change between the rods and the analyte [12–14]. The bandgap structure is determined for the proposed design prior to the insertion of defect with the help of Opti-FDTD software which employs the plane wave expansion (PWE) (see Fig. 3) shows three PBG range of normalized frequency. They are from 0.49654 to 0.78543, 0.932657 to 1.01043, and 1.29936 to 1.30245 for TM mode, and the corresponding wavelength intervals are 1273.18 nm to 2013.93 nm, 989.677 nm to 1072.20, and 767.78 nm to 769.60 nm, respectively.

The first wavelength interval is utilized for detecting purpose for the proposed structure which lies inside the third window of optical communication. As a result,

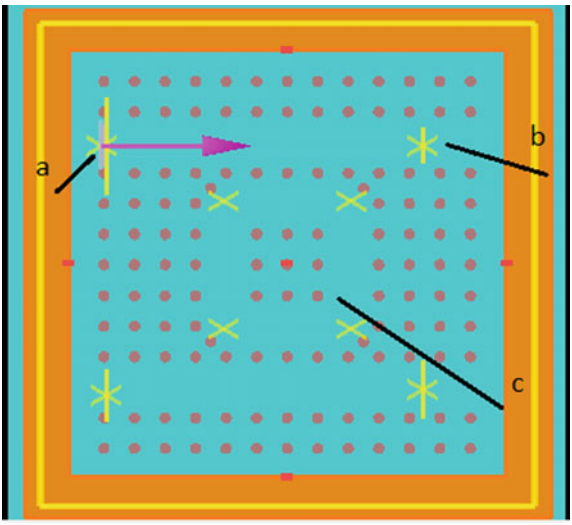


Fig. 1 Proposed sensor design

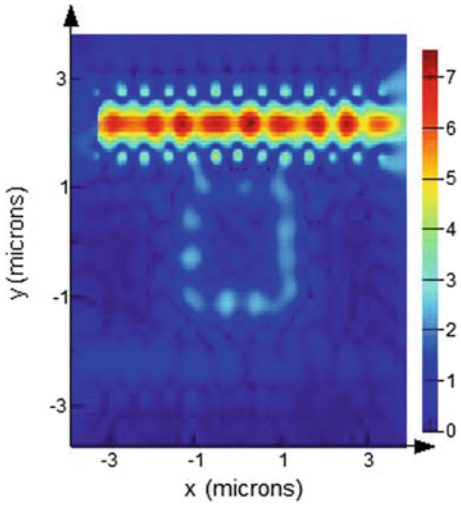


Fig. 2 Field distribution profile for proposed sensor using Lumerical FDTD

the operating wavelength for input source is selected from 1500 to 1650 nm to enable efficient sensing. The perfectly matched layer (PML) is used as a boundary surface which reduces back reflections. The substrate material is SiO_2 according to the study. The RI of material silicon rods in the suggested construction are surrounded by the RI of the analyte. The following Table 1 shows the proposed sensor design specification.

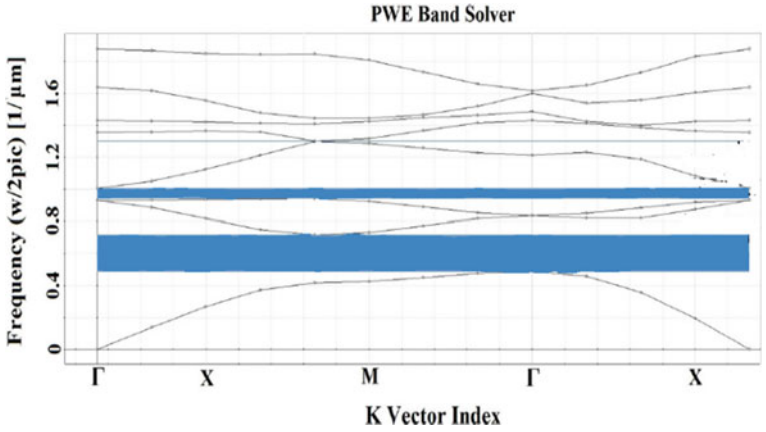


Fig. 3 Band gap structure for TM mode

Table 1 Proposed sensor design parameter

S. No	Parameters	Value
1	Photonic sensor configuration	Rods in the air configuration
2	Structure of lattice	Square structure
3	Constant (a) of the lattice	540 nm
4	Source wavelength	1480–1650 nm
5	RI of silicon	3.45
6	Radius of silicon rods	100 nm
7	Rods height	360 nm
8	Rods along the X and Y direction	13 × 13
9	Substrate material	SiO_2
10	Footprint	Around 42.5 μm^2

4 Simulation Results

The propagation of electromagnetic waves in conventional waveguides uses the total internal reflection concept. However, in photonic crystal-based structures, the waveguide is formed by utilizing line defect in the structure, and only a certain range of frequencies is permitted to propagate based on the photonic band gap generated by the periodical arrangements of two dielectric materials. The structure band gap is susceptible to modification by changing the lattice constant of the Ph.C. structure and the radius of the rods, allowing the structure to choose a certain frequency range to propagate. Because Ph.C.-based MSRRs have higher sensitivity than waveguide sensors, they are regarded as more promising structure for sensing applications. The Ansys FDTD tool is used to build and simulate the Ph.C.-based MSRR structure,

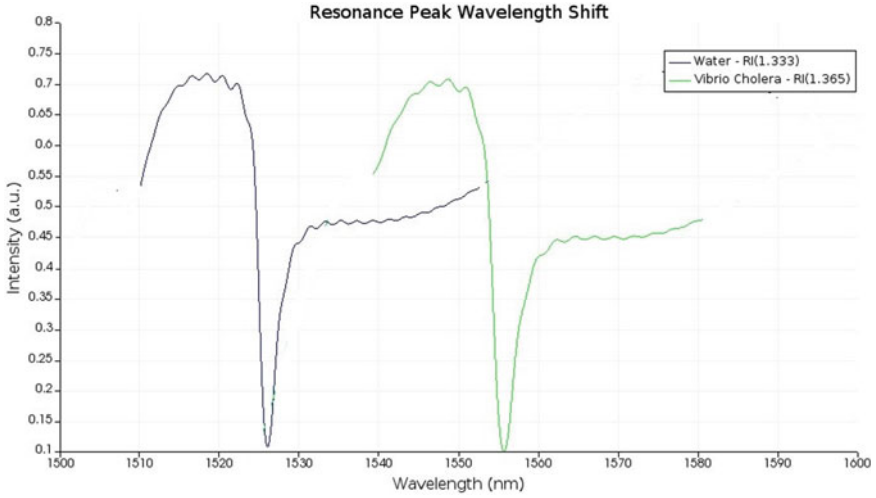


Fig. 4 Transmission graph for pure water (RI:1.333) and water with *Vibrio cholera* bacteria (1.365)

and the shift in the peak wavelength caused by the various aquatic bacteria in the water sample is evaluated using the simulation results [15].

4.1 Results for Detection of Vibrio cholera (Refractive Index: 1.365)

The output transmission graph for pure water with refractive index 1.333 and *Vibrio cholera* with refractive index of 1.365 (see Fig. 4) and the obtained results.

4.2 Results for Detection of E. coli (Refractive Index: 1.388)

The output transmission graph for pure water with refractive index 1.333 and *E. coli* with refractive index of 1.388 (see Fig. 5) and the obtained results.

4.3 Results for Detection of Shigella flexneri (Refractive Index: 1.388)

The output transmission graph for pure water with refractive index 1.333 and *Shigella flexneri* with refractive index of 1.422 (see Fig. 6) and the obtained results.

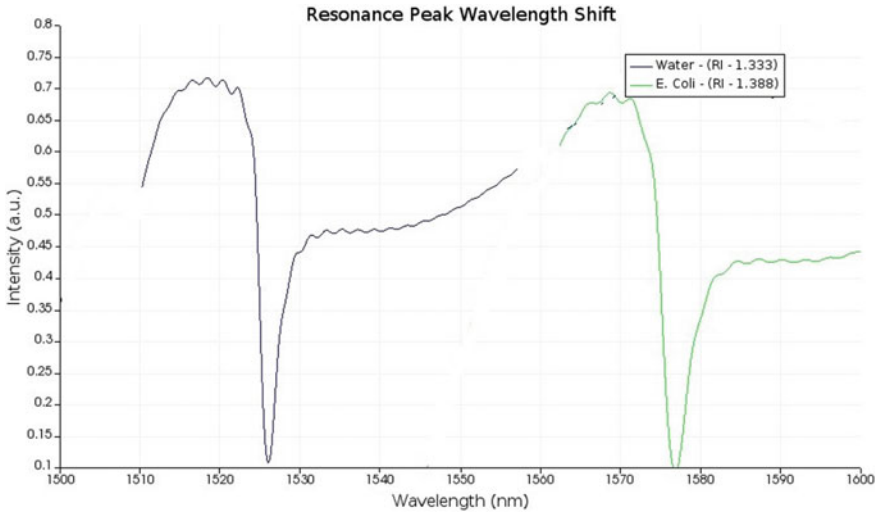


Fig. 5 Transmission graph for water (RI:1.333) and water with *E. coli* bacteria (1.388)

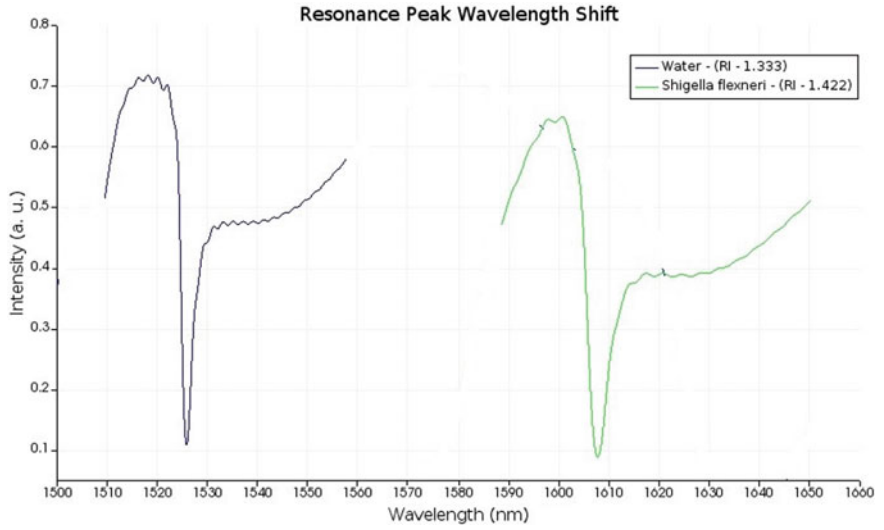


Fig. 6 Transmission graph for water (RI: 1.333) and water with *Shigella flexneri* (RI: 1.422)

4.4 Calculation of Performance Parameter

Sensitivity Calculation

The sensitivity S is defined as the ratio of the change in the peak wavelength with the change in the refractive index of the water sample due to presence of waterborne

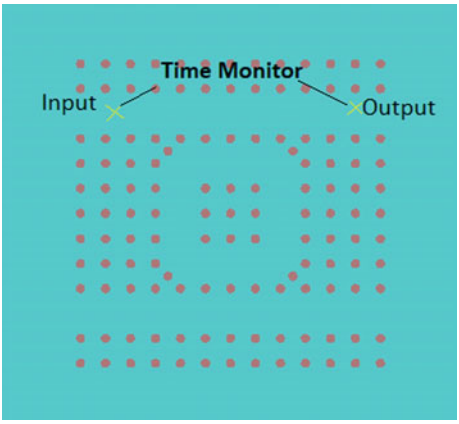


Fig. 7 Time monitor position on the proposed sensor to calculate response characteristic

bacterium present in the water given by Eq. 1 [16].

$$S = \frac{\text{Change in Resonance Peak Wavelength}}{\text{Change in the RI of Water sample}} \tag{1}$$

Sensor Response Characteristic

The sensor response characteristic was calculated using a numerical simulation approach based on FDTD. Figure depicts time monitors mounted at the input and output ports to calculate the time required by the proposed device to give output (see Fig. 7). The proposed single MSRR-based water analysis sensor seems to have a response time of around 80 fs from the characteristic graph (see Fig. 8).

Quality Factor (Q-Factor)

Q-factor is calculated in terms of peak wavelength and 3 dB bandwidth of the spectral range as given in Eq. 2.

$$\text{Q-factor} = \frac{\text{Resonance Peak Wavelength}}{\text{3 dB Bandwidth of the spectral range}} \tag{2}$$

The optimized value of Q-factor is about 646.6 for the water sample at the resonance peak wavelength of 1525.98 nm.

Figure of Merit (FOM)

Figure of merit is defined as the ratio of the sensitivity of the sensor to the 3 dB bandwidth of the spectral range. FOM is calculation is done by given Eq. 3.

$$\text{FOM} = \frac{S}{\text{3 dB bandwidth of the Spectral range}} \tag{3}$$

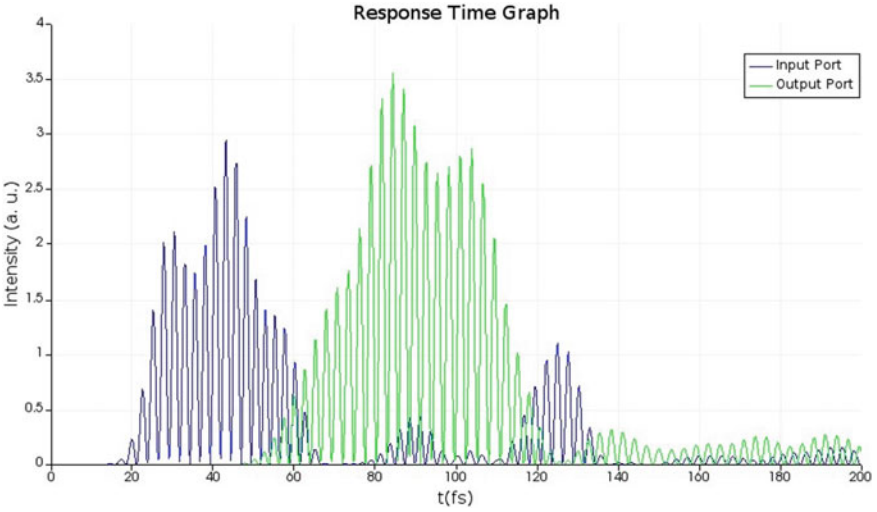


Fig. 8 Response characteristic of proposed sensor

The FOM for the proposed sensor is 392.61/RIU. Higher figure of merit of the sensor indicate a good label-free biosensor performance.

4.5 Result Observation

A photonic-based micro square ring resonator structure is designed for the detection of single waterborne bacterium present in water. Refractive index value, resonant wavelength shift, and sensitivity of normal water and water with various single waterborne are given in Table 2.

The performance parameter of proposed sensor is compared with previous published research work, and it shows that the sensor is ultra-sensitive and compact in size. The following Table 3 gives the comparison between the proposed sensor and published research work.

Table 2 Resonance peak wavelength shift and sensitivity for different waterborne bacterium

S. No	Waterborne bacterium	Refractive index	Resonance peak wavelength (nm)	Wavelength shift (nm)	Sensitivity (nm/RIU)
1	Water	1.333	1525.79	–	–
2	*Vibrio cholera*	1.365	1555.63	29.65	926.56
3	*E. Coli*	1.388	1576.82	50.84	924.36
4	*Shigella flexneri*	1.422	1607.71	81.92	920.44

Table 3 Comparison of performance parameter of proposed work with another published research work

References	Sensitivity	Proposed structure	Simulation platform used	Application of sensor
Chen et. al. [17]	95 nm/RIU	Heterostructure Ph.C. Ring Resonator	FDTD	Biosensing
Sharma et. al. [18]	343 nm/RIU	Ph.C. Ring Resonator	MIT MEEP	Detection of different Blood constituents
Chopra et. al. [19]	Not mentioned	Hexagonal Array type Ph.C. Ring Resonator	FDTD	Biosensing
Tavousi et. al. [20]	180 nm/RIU	Ph.C. Ring Resonator	FDTD	Glycalated Hb detection
Danaie and Kiani [21]	720 nm/RIU and 638 nm/RIU	Ph.C. based structure	FDTD simulation and experimental work	Cancer Cell detection
Arunkumar et. al. [22]	Not mentioned	Ph.C. Ring Resonator	FDTD	Biosensor for detection of Blood constituents
Uttara et. al. [23]	700 nm/RIU	Ph.C. based MRR	FDTD	Detection of Hb concentration in Blood
Yashaswini et. al. [24]	400 nm/RIU	Hexagonal Ring Resonator Structure	FDTD	Detection of Glucose concentrations in Blood
Proposed work	924.56 nm/RIU	Ph.C. based MSRR	FDTD	Water Analysis for detection of Waterborne Bacterium

5 Conclusion

The detection of single cell waterborne bacterium is done with the help of designed Ph.C.-based MRR and simulated using Ansys Lumerical software. The proposed sensor is designed using square lattice structure and rods in air configuration and its operating wavelength range is from 1500 to 1650 nm. The proposed sensor is providing a quality factor of 646.6, sensitivity of 926.56 nm/RIU, 924.36 nm/RIU, and 920.44 nm/RIU for the waterborne bacterium *Vibrio cholera*, *E. coli*, and *Shigella flexneri*, respectively, and FOM of approximately 392.61/RIU. The research is mainly for the design of ultra-sensitive sensor for water analysis having high speed, compact, and low cost. So, the proposed sensor performance characteristic is compared with

the published research work as shown in Table 3, and from the table we can conclude that the proposed sensor is ultra-sensitive for the detection of waterborne bacterium.

References

1. World Health Organization (WHO) (2002) The Guidelines for drinking water quality, Mibrobiological Aspects. WHO Geneva, Switzerland
2. Yablonovitch E (1993) Photonic band gap structures. Opt Soc Am 10(2)
3. Granum PE, Lund T (1997) Bacillus cereus and its water poisoning toxins. FEMS Microbiol Lett 157:223–228
4. Tuminello PS, Arakawa ET, Khare BN, Wrobel JM, Querry MR, Milham ME (1997) Optical properties of *Bacillus subtilis* spores from 0.2 to 2.5 num. Appl Opt 36(13):2818–24. https://doi.org/10.1364/ao.36.002818, PMID: 18253278
5. Sharan P, Deshmukh P, Roy SK (2013) Mapping of aqua constituents using photonic crystal. In: 2013 IEEE region 10 HT conference, Sendai, pp 320–325
6. Lavanya J, Roy SK, Sharan P (2014) Design of optical sensor for detection of brininess of water. In: 2014 IEEE global humanitarian technology conference—South Asia Satellite (GHTC-SAS), Trivandrum, pp 99–104
7. Kerimani PK, Gudagunti FD, Sharan P, Talabuttula S (2015) Photonic crystal based nano scaled sensor for the detection of agents causing typhoid in water. In: IEEE ICCSP conference, pp 1171–1175
8. Sharma V, Kalyani VL (2017) Designing four-channel nano cavities coupled photonic crystal based bio-sensor for detection in water Bactria. In: International conference on computing and communication technologies for smart nation (IC3TSN), IEEE, pp 84–88
9. Roy SK, Sharan P (2019) Design of ultra-high sensitive biosensor to detect *E. coli* in water. Int J Inf Technol
10. Nehal SA, Roy D, Devi M et al (2020) Highly sensitive lab-on-chip with deep learning AI for detection of bacteria in water. Int J Inform Technol 12:495–501. https://doi.org/10.1007/s41870-019-00363-1
11. Biswas U, Rakshit JK, Bharti GK (2020) Design of photonic crystal microring resonator based all-optical refractive-index sensor for analyzing different milk constituents. Opt Quant Electron 52(1):19
12. Joannopoulos JD, Meade R, Winn J (1995) Photonic crystals. Princeton University, Princeton, N.J.
13. Johnson SG, Joannopoulos JD (2003) Introduction to photonic crystals: Bloch's theorem, band diagrams, and gaps (but no defects). Photonic Crystal Tutorial 1–16
14. Johnson SG, Fan S, Villeneuve PR, Joannopoulos JD, Kolodziejski LA (1999) Guided modes in photonic crystal slabs. Phys Rev B 60(8):5751
15. Lumerical FDTD solutions [Online]. https://www.lumerical.com/tcad-products/fdtd/
16. Sultana J et al (2018) Terahertz detection of alcohol using a photonic crystal fiber sensor. Appl Opt 57(10):2426–2433
17. Chen Y, Wang W, Zhu Q (2014) Theoretical study on biosensing characteristics of heterostructure photonic crystal ring resonator. Optik-Int J Light Electron Opt 125(15):3931–3934
18. Sharma P, Sharan P (2015) Design of photonic crystal-based ring resonator for detection of different blood constituents. Opt commun 348:10–23
19. Chopra H, Kaler RS, Painam B (2016) Photonic crystal waveguide-based biosensor for detection of diseases. J Nanophotonics 10(3):036011
20. Tavousi A, Rakhshani MR, Mansouri-Birjandi MA (2018) High sensitivity label-free refractometer-based biosensor applicable to glycated hemoglobin detection in human blood using all-circular photonic crystal ring resonators. Opt Commun 429:166–174

21. Danaie M, Kiani B (2018) Design of a label-free photonic crystal refractive index sensor for biomedical applications. Photonics Nanostruct-Fundam Appl (2018). https://doi.org/10.1016/j.photonics.2018.06.004
22. Arunkumar R, Suaganya T, Robinson S (2019) Design and analysis of 2D photonic crystal-based biosensor to detect different blood components. Photonic Sens 9(1):69–77
23. Biswas U, Rakshit JK (2020) Detection and analysis of hemoglobin concentration in blood with the help of photonic crystal based micro ring resonator structure. Opt Quant Electron 52(10):1–16
24. Yashaswini PR, Gayathri HN, Srikanth PC (2021) Performance analysis of photonic crystal-based biosensor for the detection of bio-molecules in urine and blood. Mater Today Proc

Analysis of Energy Efficient Routing Protocol for Wireless Sensor Network in Environmental Monitoring

Rahul Pethe and Namrata Mahkalkar

Abstract Wireless sensor network is now being increasingly popular in today's world and is getting used in many kinds of fields and applications where different protocols are being used for the better communication. There is also a great requirement for the improvement of different parameters related to the different issues in this domain. This paper specifically discuss the performance related to increasing the energy efficiency of the hybrid protocol. In new hybrid protocol design which will increases life and energy efficacy of the network using the specifications of the various network QoS tools more particularly system concerns with network simulator NS2 and different parameters and issues regarding the same.

Keywords Energy efficiency · Hybrid protocol · Lifetime · QoS

1 Introduction

Wireless medium consist of various small power microdevices which are used in communication such as nodes sensors have high capability of sensing which incrassating the network performance before sending their results to the synch. Scarcity of the resources and battery power loss lines which no infrastructure and non-renewable energy supply must require the protocol which will provide efficient management of the different network resources which include following issues.

1. Consumption of power limits for nodes that use batteries and energy harvesting.
2. It has ability to handle node failures.
3. Nodes are heterogeneous.
4. Nodes are mobile.
5. It has capability to stand in strict environmental conditions.

R. Pethe
G.H.Raisoni Institute of Engineering and Technology, Nagpur, India
e-mail: rahul.pethe@raisoni.net

N. Mahkalkar (✉)
Priyadarshini College of Engineering, Nagpur, India
e-mail: namratamahakalkar@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_12

6. Scalable for huge development scale.
7. It supports cross-layer design.
8. It is simple and easy of use.

Cross-layer design has become a significant study part for the wireless communications. Traditional layer approach has presented by the three major problems. Previous layer approach cannot exchange the information between different layers due to this layer did not have the entire information. Traditional layer approach does not guarantee optimization of entire network. It didn't accept to environmental change. Traditional layer approach for the wired network is not considered and valid to the wireless level networks owed to management of environment in to the WSN [1]. So, cross-layer is measured to create the best inflection in order to expand the protocol performance. These are data rate, power effectiveness and value of service. Nodes consist of a processing unit which is having a restricted computational energy and imperfect memory. One or exceeding components of wireless sensor networks are base stations which are more calculation, energetically, and communicational. Routing based other main components are routers, which are designed to calculate, compute, and dispense routing tables.

The routing table consists of suitable route to transfer the data. It allows transmission of data without any issue to different routes. This table chooses the path which will be shorter to transfer the data. The short path selection helps in fast data transmission as well as it reduces the probability of loss of data during transmission.

2 Background

A wireless sensor network (WSN) consists of large number of low power, low cost, and tiny communication devices, called sensors. Like nodes (i.e., computers, laptops, etc.) in traditional wireless networks such as mobile ad hoc networks, sensors have energy, storage, processing, and communication capabilities. Also, sensors have a sensing capability by which they sense phenomena and perform in-network processing on the sensed data before sending their results to a central gathering node, called the sink. WSNs can be used in a variety of monitoring, control, and surveillance applications [2]. Particularly, the sensors possess several scarce resources, with battery power (or energy) being the most critical one. One way to extend the lifetime of a WSN is through load balancing so that all the sensors deplete their energy as slowly and uniformly as possible. Also, the behavior of the sink has an impact on the network lifetime. Indeed, sensors in the proximity of a static sink act as the traffic hot spots have significantly reduced lifetime than all other sensors in the network. Those sensors nearer a static sink would suffer from a severe depletion of their battery power, which may result in possible network disconnection and disruption of the data from reaching the sink. Sariga Arjunan et al. Elsevier, (2017). A survey on unequal clustering protocols in wireless sensor networks; a comprehensive survey of various unequal clustering approaches with their objectives, characteristics, etc.,

is presented. Also, the classifications of unequal clustering approaches are made and compared based on various cluster properties, cluster head (CH) properties, and clustering process.

3 Previous Work Done

Most of the monitoring system evolved from the previous few years used the wireless sensor network technology for providing the facilities regarding the monitoring of plant health such a system are also using the Radio + CPU + sensing technology for the continuous monitoring [1].

Such systems consist of the autonomous devices called as sensors which monitor the different environmental and physical conditions. The environmental conditions may differ from location to location and area [2] specific constraints are also involved.

In recent technology also there is continuous need of the wireless sensor network as an important technology in the field of agricultural domain. The another advantage which must be noted while using the IOT as latest trend in the application of the agricultural monitoring of the crop plant is the speed and accuracy of data which is considered as one of the significant parameter in this process.

4 Existing Methodologies

Trends in area of research are in the field of different domain and field:

i. Agriculture domain
ii. Environmental monitoring
iii. Urban terrain
iv. Trackingsssss and civil structure monitoring.

International Trends:

i. Health care systems
ii. Transportation and logistics
iii. Security and surveillance.

To generate minimum cost and also used the smallest sensor nodes for better efficiency is the main objective of the such a system. Counting of small companies which produce wireless sensor networks hardware is increasing as compared to counting in 1970s. Intrinsic to the sensor networks implementation is in use of the least amount of power methods for the two way radio communication and data attainment. This gateway plays an overpass between other networks with wireless sensor networks. Now data is able to be processed as well as stored by devices with added resources, for example, in any server located at a distance.

The scarcest causes of sensor networks that are wireless nodes are energy. Survival time of networks that are wireless sensor based is also resolute by it. Wireless sensor networks could also be deployed in a large number in the several environments, which consists of hostile as well as remote areas, where ad hoc communications are the main component.

The operating system complexity of the WSN knobs is characteristically fewer than the common purpose working schemes. Due to the two reasons, these are the extra strong resembled implanted systems. The first motive is that the WSNs have been organized with unique client in to the brain, besides the common platform. The other one is needed of low cost as well as low power. It leads to that the mainly WSN knobs should also have the low-control consuming micro-controllers which ensure that the essential memory mechanisms moreover pointless or the most costly for the implementation.

5 Analysis and Discussion

First, we analyze the AODV protocol and find out the energy efficient and life time maximization based on quality of service parameter. Then we form clusters of the nodes assign cluster head virtual cluster head and calculate the energy based on quality of service Qos metrics such as energy packet delivery ratio throughput, and dropping of packet. Communication and packet transfer between different nodes by using virtual cluster head in single cluster and multiple clusters also. Performance analysis of all the different techniques is also done regarding the different parameters. The method also discusses comparative analysis with other protocols of warless sensor networks.

Life time is measured using the throughput of the network. Wireless sensor network different protocols and their comparison [3] AODV and DSR study of these protocols quality of service parameters. The analysis is also made for cluster forming process. Optimization of energy and increasing the life span of the network are the most significant issues in maximization process. Routing energy efficiency for better performance and security must also be considered. For the proper analysis of the monitoring system, there are certain parameters are very important and those needs to be studied and addressed very carefully. Another important aspect is the use of transmission protocol to transmit the essential information in various forms such as image. The image coming from different devices particularly camera used in the system for monitoring of various constraints. The system must suggest the correct alternative for the proper operation and must maintain the various environmental components like pressure, temperature and humidity essential for the growth of the particular plants [4].

The success of any system used for monitoring depends on many components, but the finally it has to deal with the crucial aspects of the network transmission time and speed. There are many challenges that need to be addressed in the concerning to the monitoring of the wireless network system.

6 Proposed Methodology

The method for the proposed system includes the different parameters regarding the improvement of the performance of the complete system for the wireless network using the specific hybrid protocol.

Simulation parameters tables indicate the different regarding the wireless network and other related issues like terrain size 1200 m^2 simulation time 30,000 millisecond number of nodes 30 or 40 and transmission range 250–280 m. Another important aspect of this research and about the proposed system is that there is proper transmission of the data in the system. As it is being continuously being generated in the system with the help for the different sensors like temperature and humidity sensors these available data has to be first converted in the proper form and must be communicated over the different modules of the system in order to take the proper decision by the monitoring system (Fig. 1).

The algorithms works as follows,

1. Take input from temperature, humidity and moisture sensors.
2. Convert it into digital format using inbuilt analog to digital convertor placed on Adriano board.
3. Transfer data to server using Zig-Bee wireless communication protocol and store in database.
4. Find Euclidean distance at each point is calculated.
5. Communication and packet transfer between different nodes without forming cluster by using protocol and Euclidean's distance algorithm **DRWSN**.

The above model helps to take the decision in the system according to different sensor values. A hidden Marko model is a statistical model that can be used to describe the evaluation of observable events that depend on the internal factors which are not directly observable.

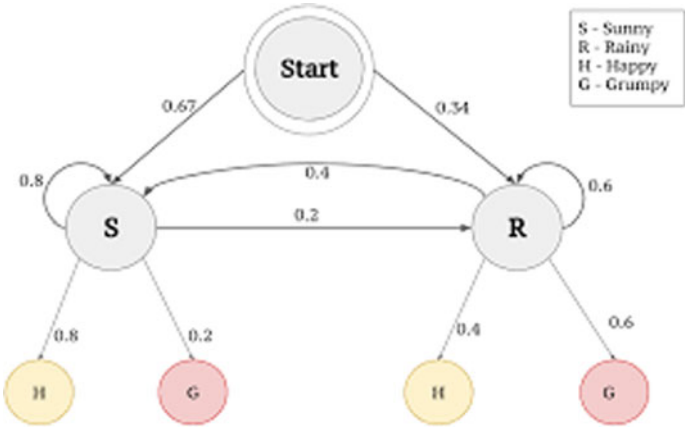


Fig. 1 Proposed hidden Markov model

7 Outcome and Results

This will include the how to deploy the nodes and packet drop simulation method and other simulation parameters like number of rounds. How it is enhances the network energy and throughput and other significant parameters in the wireless sensor network.

The system can be installed on various platform for the accessing the various kinds of environment variables based on the need of the application. In more particular, it is likely be implemented in the agriculture domain for improving the performance issues. Since then, the continuous evaluation and assessment of the various components can be done according to the various demands of the applications, and further, the need can be converted into the part of the facility which is very essential for the implementation of the other aspects and dimensions regarding the application domain. This proposed system can be efficiently and conveniently used for the implementation and design of the new hardware model which will be more useful in the various domains. Further, this proposed work is very much useful in many applications and also supportive in all environmental situations. So it can find effective results for yielding high-quality crop plants in different weather conditions.

Within this system the number of a live node greater as compared to dead nodes. The network lifetime can also be enhanced using the method of multi-hop routing between the cluster heads and sink nodes. Such kind of system is very useful and essential for the various levels of the user. The user may be the famer or it may be the merchant or may be the businessmen. It is expected that the system once installed on the field will perform considering the various parameters in to the consideration.

The analysis is taken from the various kinds of the events and may produce the different features depending on the scenario for the various inputs and the performance of the system can be increased by increasing the accuracy of the system. Different methods discussed in earlier sections are compared in different scenarios. It can also be shown by making comparative analysis of both the methods. From this, it can be easily be conclude that the performance and accuracy of the proposed system is higher than statistical methods mentioned earlier.

8 Conclusion

The mechanism discussed in this paper is all about increasing the performance of the protocol which is used in implementing the system which increases not only the efficiency and but also reduces the energy consumption of the overall network.

Such a system can be implemented in various domain and fields such as agriculture where the actual sensing of the different soil and agricultural parameter such as moisture, wind pressure, heat, and other environmental parameters has to be accurately being measured according to the favorable situation of the plant crop.

Many such fields will be getting benefitted by using this model and can improve the performance and output by implementing the minimum energy requirement model.

9 Future Scope

It is estimated that work in this domain and continuous improvement will definitely results in a several applications of the discussed model. These methods and implementations may increase the performance and effectiveness of the previous designs.

References

1. Agrawal N, Singhai J, Agarwal DK (2017) Grape leaf disease detection and classification using multi-class support vector machine. In: International conference on recent innovations in signal processing and embedded systems (RISE). IEEE
2. Dasgupta SR et al (2020) Detection of diseases in potato leaves using transfer learning. Comput Intell Pattern Recogn 675–684. Springer , Singapore
3. Dhingra G, Kumar V, Joshi HD (2018) Study of digital image processing techniques for leaf disease detection and classification. Multimedia Tools Appl 77:1519951–20000
4. Ferentinos, Konstantinos (2018) Deep learning models for plant disease detection and diagnosis. Comput Electron Agric 145:311–318
5. Hlaing CS, Zaw SMM (2018) Tomato plant diseases classification using statistical texture feature and color feature. In: IEEE/ACIS 17th international conference on computer and information science (ICIS). IEEE
6. Iqbal Z et al (2018) An automated detection and classification of citrus plant diseases using image processing techniques: a review. Comput Electron Agric 153
7. Islam M et al (2017) Detection of potato diseases using image segmentation and multiclass support vector machine. In: IEEE 30th Canadian conference on electrical and computer engineering (CCECE). IEEE
8. Jiang P et al (2019) Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access 7
9. Khan MA et al (2019) An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection. IEEE Access
10. Kour VP, Arora S (2019) Particle swarm optimization based support vector machine (P-SVM) for the segmentation and classification of plants. IEEE Access 7
11. Kamavisdar P, Saluja S, Agrawal S (2013) A survey on image classification approaches and techniques. Int J Adv Res Comput Commun Eng
12. Liu B et al (2018) Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry10(1)
13. Lin Z et al (2019) A unified matrix-based convolutional neural network for fine-grained image classification of wheat leaf diseases. IEEE Access 7
14. Ma J et al (2018) A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. Comput Electron Agric 54
15. Ogundile OO, Balogun MB, Ijiga OE, Falayi EO (2019) Energy-balanced and energy-efficient clustering routing protocol for wireless sensor networks. IET Commun 13(10):1449–1457
16. Elsmany EFA, Omar MA, Wan TC, Altahir AA (2019) EESRA: energy efficient scalable routing algorithm for wireless sensor networks. IEEE Access 7:96974–96983

A Virtual Tutor to Enhance the Solving Skills of School Children Using Performance Evaluation and Navigation System

Shanta Sodur, Harmeet Singh, Rohan Pol, and Mohit Kale

Abstract The proposed *AI-based Performance Analysis and Navigation System* (AI-PANS) aims at the development of self-learning app for the students of Grade 1–10 students studying in the Maharashtra (India) State Board and also for the teachers. The proposed AI-PANS is designed for: – Database Generation, to ease the process of making new questions. – Recommend question paper based on the student's current and past performance in every iteration. Modify/redefine the *Degree Of Difficulty* (DOD) based on the time taken and correctness of the attempted question by a large set of students. – Navigate the student across topics based on his/her performance and in turn help him/her in clearing the concepts, improve the problem-solving skills, build confidence, retain interest, and prepare him/her for competitive exams. – Hand-holding in the form of curated solutions in the form of text, images, audio, etc., and summary of each topic. In the proposed system, we also proposed a novel mechanism to redefine the DOD assigned to every question in the beginning.

Keywords Clustering · Database generation · Education · Performance analysis · Question recommendation · Test papers

S. Sodur (✉) · H. Singh · R. Pol · M. Kale
Vivekanand Education Society's Institute of Technology, Mumbai, Maharashtra, India
e-mail: shanta.sondur@ves.ac.in

H. Singh
e-mail: 2018.harmeet.kathoda@ves.ac.in

R. Pol
e-mail: 2018.rohan.pol@ves.ac.in

M. Kale
e-mail: 2016.mohit.kale@ves.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_13

1 Introduction

Mathematics have long held a high level of respect among their academic peers. Widespread fear of mathematics appears in many forms. If we are able to help a student in understanding the concept and develop interest in the subject by posing the problems which are simple and based on day to day activities along with the solution which is simple in words, correct and meaningful without any ambiguity, then the student gains interest and willingness in solving the next problem on his own. If the system is able to monitor his/her performance, then we can take the student either forward to the next level or else backward to the basic level (in a particular section).

2 Objectives

- To make mathematics easy to understand for students from basic to complex concepts.
- To create an application which will give *Multiple-Choice Questions* (MCQ) format tests to the students on various mathematical concepts, provide them with simple solution, and analyze their answers based on correctness and degree of difficulty.
- Free app for the school students of grade 5–10 with bilingual (English and Marathi) features for students of rural areas.
- To improve students' domain-general problem-solving skills, students' test-taking abilities, such as their time management.
- To create a recommendation system which will navigate the student to appropriate question tests according to his ability.
- To provide performance analysis at student level, class level, and school level.
- To help students and teachers to develop an interest and remove mathematics anxiety.

Machine learning and artificial intelligence (Supervised and unsupervised) are the primary domains for this project as they will be used for algorithm development, performance analysis, and students' navigation.

The scope of the project includes:

- Database Generation
- Question paper Recommendation
- Performance Evaluation
- Redefining the Degree Of Difficulty (DOD) and time of the Question
- Navigation

3 Literature Survey

3.1 Existing Applications

- **Khan Academy**: Khan Academy [1] It is an Interactive platform for students and teachers with over 4000 videos on educational topics like which includes maths.
- **Photomath**: PhotoMath [2] This is artificial intelligence-based application which uses 2 types of algorithms, namely *Optical character recognition* (OCR) and equation solving programs. User needs to click image of the written math problem, and application returns the solution.
- **Socrative Student**: Socrative Student [3] This is an application to conduct quizzes, surveys, etc.

3.2 Research Papers

Student Academic Performance Monitoring and Evaluation Using Data Mining Techniques suggests usage of data mining technique helps students expand their academic result [4].

Dynamic Question Paper Template Generation Using Bi-proportional Scaling Method elaborates on the need of constant evaluation to check progress of various cognitive skills under a subject at different stages of learning [5].

A Web-based testing system with dynamic question generation signifies the rising utilization of electronic media for education [6]

Although there has been extensive research in the field of Edtech, one significant limitation faced is the generation of questions. This often results in additional stress on the teachers. Instead of the manual method followed by a large number of applications, we decide to automate this process. Another limitation that stands out is the non-changing difficulty level of questions and the often overlooked performance of the student which results in a static system for the students to interact with.

Hence, we come up with the proposed application which aims to solve the shortcomings mentioned above.

4 Proposed Methodology

The proposed application is divided into five major parts. First part is database generation which generates the questions in an automated manner; second is question recommender for personalized recommendations; third is performance analysis responsible for analyzing the student's performance; then, fourth is the naviagtion module followed by DOD clustering module as the fifth part which dynamically changes the question's difficulty level based on relevant past data (Fig. 1).

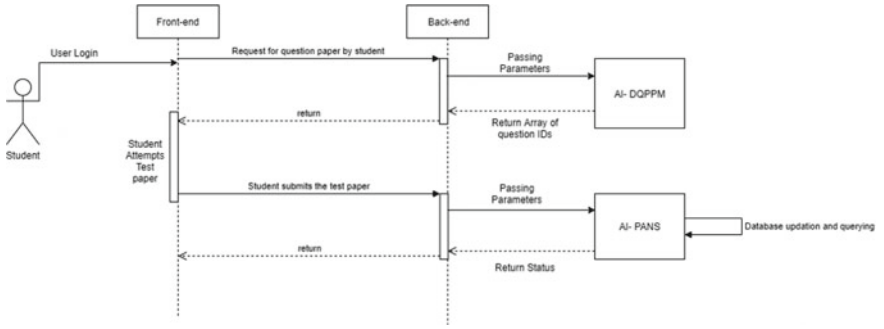


Fig. 1 Overview of the application

The above image shows the two major modules under the application. The *AI-based Dynamic Model for Question Paper Generation and Performance Monitoring* (AI-DQPPM) module includes the question recommender and navigation module. The second module is AI-PANS which consists of the performance analyzer and the DOD clustering sub-module.

4.1 Database Generation

This module consists of generating a database of varieties of problems under algebra vertical (Std V–X) using a generic problem statement. This database is generated through an algorithm designed for a generic problem with all sorts of possible variations (along with solutions), and also with different degrees of difficulties which will help the student of any category (Excellent,above average,average, below average) to clear his/her concepts from basic to advanced level. These variations include the questions of MCQ type (single correct, multiple correct, true/false, match the pair, and fill in the blanks). The questions and solutions are in the form of text, images, videos, and audios.

4.1.1 List of Sections for Algebra Vertical

- 0301 Equation
- 0302 - /, Using letter in place of a number—variable/unknown, solution to equation in variables
- 0303 Mathematical Expressions
- 0304 Algebraic Expressions

4.2 Question Recommender

Recommends personalized questions to the student based on his performance on individual questions, easiest question from the topic, and questions from students having similar performance.

This component contains three standalone modules that each recommend questions based on a different criteria.

- **Popularity-Based Recommendation**: This module queries the database and picks out the easiest, i.e., questions that have a high chance of being ticked correct by the student from that particular module. This module is useful when the student has just registered and hasn't yet attempted a lot of questions.
- **User based recommendation**: This module uses the student's performance to pick out questions according to it, with the goal of increasing his/her performance after each iteration of a question paper. This module helps the student by recommending them the questions they might not be doing well in.
- **Collaborative Filtering**: The main task of this module is to recommend newer questions to the student which he hasn't attempted yet. It achieves this by finding out other similar student to him/her in performance and then recommending the questions they have performed well in.

4.3 Performance Analyzer

Measures the performance of the student based on correctness, time taken, and DOD of the questions considering past history. For every correct attempt, the performance of the student is incremented proportionally to the difficulty level the question had, and the time he/she took to answer the question. The same results in a decrement in performance level if the attempt goes wrong. Also, the decrement in performance is directly proportional to the time taken to solve the question.

This parameter will be used to calculate the student's performance for each variation the student attempts and use it as a marker to judge his current proficiency level in the corresponding topic.

The above image also shows the DOD clusterer module which will be discussed further in this section.

4.4 Navigation

The main function of this module is to determine whether the student should be promoted or demoted to the next or previous module based on his performance on the current topic. If the student's performance surpasses the median value of 4 in the

current topic, then this module will move him to the next respective topic. Similarly, when his/her median performance in the current topic goes below 1, the module will demote him to the previous respective module as the student might need some more practice to attempt the questions of the current topic.

4.5 DOD Clustering Module

This module clusters all the variations in 5°C of difficulty (DOD) ranging from 1 to 5. The clustering algorithm is implemented on 2 levels. In the first level, all the variations are are clustered using the *Corrections Ratio or Success Ratio*, defined as

$$\text{Correction Ratio} = \frac{\text{Correctly Attempted}}{\text{Attempted}} \tag{1}$$

where CorrectlyAttempted is the number of times the question was attempted correctly and *Attempted* is the number of time the question is attempted.

In the second level, the variations are clustered according to the *time taken* by the student; hence, the clusters created are called as *time based clusters or T-clusters*. The *T-clusters* range from 1 to 5. If a variation has *T-cluster* of 1, then the time taken by students to solve that variation is very less than expected. Similarly, if *T-cluster* of a variation is 5, then time taken by students to solve that variation is very high. So, in conclusion, the variations with *T-cluster* of 1 are considered easier, and the variations with *T-cluster* of 5 are considered difficult. The proportional difference between expected time and actual time taken is used for *Time based clustering*, and it is calculated using the formula given below

$$\text{Expected Time} = \text{Correctly Attempted} * \text{Time Assigned} \tag{2}$$

where Time Assigned, the time expected by the professionals that the student should take to successfully solve the question.

$$\text{Proportional Difference} = \frac{\text{Time Taken} - \text{Exptected Time}}{\text{Expected Time}} \tag{3}$$

where Time Taken is the cumulative time taken by all the students to attempt the question correctly.

After dividing variations in 5 clusters using both the methods, the final difficulty level is assigned using swapping method. For example, If a variation with id 101 has correction ratio-based cluster as 3 and it has T-cluster as 5, then that variation will be assigned DOD of level 4. Whereas if the same variation had the T-cluster of 1, then the DOD 2 should be assigned to it.

	A	B	C	D	E	F
	Sr. No	uestion Ty	nswer Typ	pic Numb	Question (Text Only)	Correct Answer 1
1		text	1	030101	Which of the following stateme	$22z - 27 = 22$
2		text	1	030101	Which of the following stateme	$4a + 6 = 25$
3		text	1	030101	Which of the following stateme	$14a - 18 = 30$
4		text	1	030101	Which of the following stateme	$20a + 12 = 25$
5		text	1	030101	Which of the following stateme	$8y + 10 = 18$
6		text	1	030101	Which of the following stateme	$22y + 17 = 27$
7		text	1	030101	Which of the following stateme	$30b + 24 = 20$
8		text	1	030101	Which of the following stateme	$6c + 11 = 29$
9		text	1	030101	Which of the following stateme	$16y - 13 = 29$
10		text	1	030101	Which of the following stateme	$9b + 11 = 15$

Fig. 2 Snapshot of Excel sheet generated in module 030101

5 Implementation

This section discusses the implementation part of our proposed system. First, we show the outcome of coep-package developed to standardize the question generation process. Next, we show the integration of the proposed system with a Web site through APIs along with the performance analysis and navigation part.

5.1 CSV Generation

See Fig. 2.

5.2 Navigation

This section shows how the student will be navigated across the sections/subsections in the current topic (Fig. 3).

5.3 DOD Clustering

See Fig. 4.

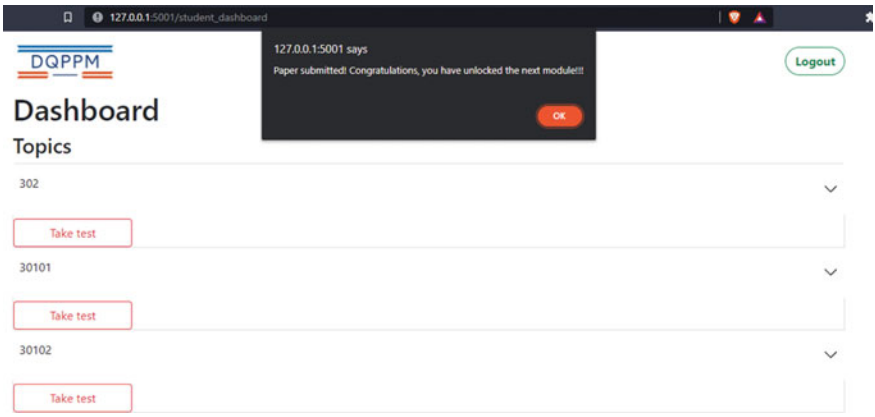


Fig. 3 Students being promoted after she/he reaches the performance of 4.5

CorrectlyAttempted	Attempted	Total_Time	clusters	time_clusters	suggested_time_in_seconds	suggested_dods
993	1520	179371	4	1	54	3
1264	1798	222855	2	2	73	2
1142	1703	204393	1	3	78	1
1082	1609	192064	3	3	80	3
991	1527	180685	4	3	78	4
1008	1532	180406	4	4	86	4
1095	1641	188560	3	3	79	3

Fig. 4 Snapshot after executing DOD clustering algorithm

6 Results

We present our results simulated through an algorithm which attempts the question paper randomly for a set of 30 iterations See Fig. 5.

Finally, we infer that we have simulated the solving process of the student anticipating the student may follow this pattern.

However, the exact results would be available only when the question paper is rendered to the students of respective grade.

7 Conclusion

We conclude that, we have successfully implemented the proposed algorithm and tested the performance through simulated results against the expected ones.

From the results, we can see that we have automated the CSV generation with multiple questions populated based on the requirement, through our package *coep-*

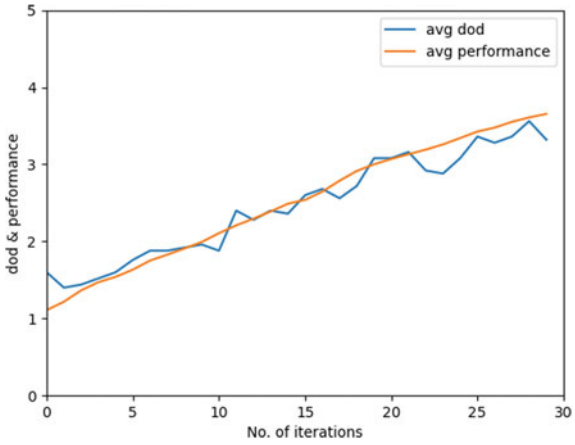


Fig. 5 Graph of Performance against No. of iterations for 30 question papers

package. We also devised the technique for DOD clustering, and finally, we have demonstrated the navigation of student based on his/her performance.

Acknowledgements We thank the Department of Instrumentation and Control Engineering of College of Engineering (COEP) for giving us an opportunity to collaborate with them to work on the Virtual Laboratory Project of the Ministry of Education, Govt. of India, under the "Evaluation framework for virtual labs." We also thank Prof S.D. Agashe and his team and Mr. Dhananjay Thosar for their invaluable guidance and support in implementing this prototype.

References

1. Khan Academy. Accessed 29 Mar 2022. https://www.khanacademy.org
2. PhotoMath. Accessed 29 Mar 2022. https://photomath.com/en/
3. Socrative Student. Accessed 29 Mar 2022. https://www.socrative.com/
4. Ogor EN (2007) Student academic performance monitoring and evaluation using data mining techniques. In: Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007), pp 354–359. https://doi.org/10.1109/CERMA.2007.4367712
5. Paul DV, Pawar JD (2013) Dynamic question paper template generation using Bi-proportional scaling method. In: 2013 IEEE fifth international conference on technology for education (t4e 2013), pp 80–83. https://doi.org/10.1109/T4E.2013.28
6. McGough J, Mortensen J, Johnson J, Fadali S (2001) A web-based testing system with dynamic question generation. In: 31st annual frontiers in education conference. Impact on Engineering and Science Education. Conference Proceedings (Cat. No.01CH37193), pp S3C-23. https://doi.org/10.1109/FIE.2001.964056

A Survey of Learning Techniques for Detecting DDOS Assaults

K. Jeevan Pradeep and Pragnyaban Mishra

Abstract The distributed denial-of-service (DDOS) exploit is one of the most catastrophic assaults on the Internet, disrupting the performance of critical administrations offered by numerous organizations. These attacks have become increasingly complicated, and their number has been steadily increasing, making it harder to detect and respond to such assaults As a result, a sharp security system (IDS) is necessary to detect and control any unexpected system traffic behavior. In a DDOS Assaults, the intruder delivers a stream of packets to the server while exploiting known or unknown flaws and vulnerabilities.

Keywords DDOS assaults · Network security · Decision tree · Naïve Bayes · SVM · Neural network · Fuzzy logic · Learning techniques

1 Introduction

Communication infrastructure and information assurance play critical roles in both social and economic growth, as well as in our daily lives. Because of the fast growth of world wide web-related networking and communication networks, knowledge management are becoming increasingly exposed to a variety of cyberattacks. Communication network and system attacks are becoming a security concern. Cyber security assaults are on the rise in general. As a result, a comprehensive security system for detecting security breaches is essential to defend the network from all forms of assaults.

An intruder attempt is defined as an unlawful attempt or threat to (i) acquire information, (ii) modify or alter the content, or (iii) disable the system. As an example,

K. Jeevan Pradeep (✉) · P. Mishra
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
e-mail: jeevanpradeep33@gmail.com

P. Mishra
e-mail: pragnyaban@kluniversity.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_14

A. Malware exploits of other network nodes by interfering with routine operations by corrupting software with a fault (virus) or polluting the network with a bug. As a result, network traffic is slowed.
B. A denial-of-service (DoS) attack is one in which the attacker intends to make a system or a network resource inaccessible to legitimate users.

A distributed denial-of-service (DDoS) attack is one that uses several machines in a distributed method to target a victim. This research looked into the usage of machine learning algorithms to detect DDoS assaults. Every cyber-attack leaves a digital footprint. An IDS uses a signature, which is a collection of criteria, to detect harmful intrusive behavior in the network, such as DDoS attacks. Among the approaches for identifying signatures are:

- An attempt to connect from a reserved IP address was made. An IP header connection might be easily recognized by examining the source address information.
- In a packet, an illegal TCP flag combination was used. Positive and negative flag mixes can be detected by comparing the flag set in a TCP header to known positive and negative flag mixes.
- An email server has been infected with a virus. This may be discovered by inspecting the subject lines of every email sent to the subject of a virus-infected email.
- DOS attack on a POP3 server caused by sending the same command several times. One signature will be kept to keep track of how many times the same command has been issued and to alert if the number of times the same command has been issued exceeds a specific threshold.
- Attack the FTP server's file system by supplying directories and file commands without even logging in. A state tracking system may be extended to track FTP traffic to ensure proper sign-in. This might alert you if you issued any of the instructions before the user was authorized.

Several real intrusion detection systems rely on human research to distinguish between intrusive and non-invasive assaults. Because human intervention is required to build, debug, and deploy the extension on the studied datasets, finding and creating a new signature for an attack may take a long time or several hours.

Because of the catastrophic consequences of a compromised system on commercial and personal networks, intrusion systems have become a primary study area for researchers, cybersecurity administrators, and network administrators. An intrusion detection system (IDS) can identify several forms of harmful network activity and hostile network incursions, but a normal firewall cannot. In general, intrusion detection systems are divided into two types: Signature-based IDSs and anomaly-based IDSs.

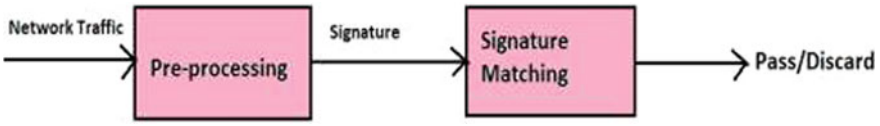


Fig. 1 Signature-based detection

1.1 Signature-Based Detection

Signature-based approaches rely on prior observed signatures to tally the signatures that are recorded in a database. This database contains a collection of signatures connected with prior assaults. Signature of IDs with some precise information in computation and preparation so that it does not check for every activity or network traffic on the environment that is being monitored. The signature-based technique is simple to use since it does not need learning about the environment; instead, known signatures are previously recorded in the database. The signature-based technique is effective against known attacks, but it cannot detect new attacks until it is activated with fresh signatures.

Signature-based intrusion detection systems are extremely difficult to evade since they are based on known attacks and require a new signature to be used before they can be recognized as new assaults. Signature-based approaches are easy to alter and enhance because their performance is determined by the signature or rules used.

Figure 1 depicts the design of the signature-based method. The network traffic is pre-processed in this architecture to obtain the desired properties and to find the signature. The signature is then matched to the action signature and validated against the signature database. If the signatures match, an alert is triggered; if no signatures match, nothing happens. Many businesses utilize this form of intrusion detection system to identify known threats financially and with few false positive failures.

1.2 Anomaly-Based Detection

Anomaly-based detection is also known as "Behavior based detection" that's because the models are based on network behavior, and most computer systems issue an alert message when there is a normal detection in the behavior. This approach identifies undesirable traffic and is best suited for doing research on network hardware. Figure 2 depicts the overall architecture, which incorporates anomaly-based approaches. This design contains a pre-processing function that collects data and builds a pattern for the connection; if it deviates from regular behavior, an alarm message is generated.

This paper explored several machine learning techniques for detecting DDOS attacks. The following is the rest of the paper: Sect. 2 provides an overview of the literature review on machine learning approaches utilized in DDOS detection. Section 3 compares several ML algorithms for detecting DDOS assaults, as well

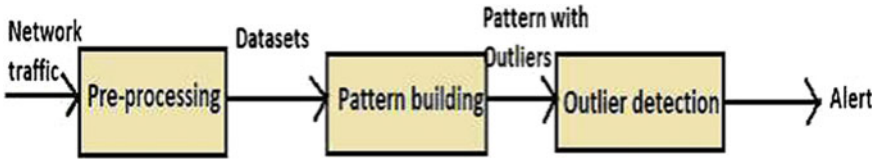


Fig. 2 Anomaly-based detection

as their benefits and drawbacks. Section 4 brings the study of machine learning techniques to a close.

2 Literature Survey

This section includes a literature overview of ML techniques used in the detection of DDOS assaults, such as decision tree, Naive Bayes, artificial neural networks, support vector machine, and fuzzy logic, as well as the research that has been conducted on them.

2.1 Decision Tree

Many researchers have been published on decision tree prediction models to identify DDOS attacks. Wu et al. [6] constructed a DDoS Detection system using a decision tree method, and the system uses a traffic-flow pattern-matching technique to trace back the attacker's location when an attack is detected. A C.45 classifier is used to detect DDOS assaults. The author of [8] discovers a method for efficiently detecting DDOS attacks. Many ML approaches take longer to identify an attack or yield lower accuracy. In [8], the C4.5 methodology is used, which has poorer accuracy and takes longer to generate the decision tree; however, the C5.0 method has been proved to be more efficient since it consumes less time and memory than the C4.5 technique. Bujlow, Khafhali, and Saadi [9] focused on network traffic categorization and found that the C5.0 algorithm outperformed the C4.5 method. Another research utilized ID3, C4.5, and C5.0 to construct a better decision tree with less error pruning and feature selection [10]. C5.0 outperformed in terms of accuracy and memory use, according to the results.

2.2 Support Vector Machines (SVM)

It is the most extensively used and popular strategy for machine learning assignments. In 2010, Das et al. [12] conducted an attempt to identify DDOS assaults using RTS and SVM. The RST preprocessed the network packet data that was first acquired. The SVM model is given the feature set chosen by the RST to train and test. When compared to principal component analysis (PCA), RST and SMV may reduce false positives and improve accuracy.

2.3 Naive Bayes

This machine learning algorithm is a basic probabilistic classifier [13]. Carl Livadas [16] employed Ml methods to identify commands and manage IRC-based botnet traffic. By comparing the performance of J48, Nave Bayes, and Bayesian networks, the author distinguishes between IRC and non-IRC traffic. The author of paper [16] discovers the qualities that provide greater accuracy. This classifier produces low false negative (2.49%) and false positive (15.04%) rates for real-life IRC/non-IRC flows, as well as low false negative (7.89%) rates for botnets tested on IRC flows, demonstrating naive Bayes to be an efficient classifier. In [14], Bains et al. suggested a layered hierarchical strategy for attack detection accuracy.

2.4 Neural Networks

Neural networks incorporate processing elements to turn a collection of inputs into a set of outputs that function in a manner comparable to organic nerve systems such as the human brain. Gavrilis et al. [15] employed a "RBF-NN detector" with nine packet parameters and computed the frequencies of these parameters. RBF-NN categorizes traffic as either normal or assault based on the expected frequencies. Distributed time delay neural network (DTDNN) has a high likelihood of detecting assaults with more precision. DTDNN performs data classification with quick conversion rates and great speed.

2.4.1 Fuzzy Logic

Anomaly detection mostly use fuzzy approaches. In IEEE 802.15.4, the author Vladimir [17] suggested a DDoS detection and prediction approach based on fuzzy logic. The fuzzy-based detection and prediction system (FBDPS) assisted in the identification of DDODS assaults by analyzing the energy usage of sensor nodes. The unusual energy usage of the node identifies it as a hostile attacker.

3 Comparision of Machine Learning Techniques

Despite the fact that machine learning-based algorithms are employed to identify intrusion in order to attain a high detection rate, they have their own set of advantages and drawbacks.

3.1 Fuzzy Logic

In [7], Yusof et al. reported that fuzzy c-means clustering outperforms other classifiers in categorization. Fuzzy c-means were shown to be faster than other machine learning methods. Fuzzy logic relies on reasoning that is approximate rather than exact. Suresh [4] used ML methods to construct and analyze fuzzy c-means clustering on DDOS assaults and obtained better categorization than previous solutions. Identification of reduced, relevant rule subsets is a tough process. In [5], Manjula and Anitha assessed ML approaches for detecting DDOS attacks using the CAIDA dataset, which is based on chi-square and information gaining ranking for the selected characteristics. The results demonstrate that fuzzy-c means better classification and is faster than other algorithms. Fuzzy logic works well against port scanning and probes.

3.2 Neural Networks

Jie-Hao et al. [2] utilized artificial neural networks to identify DDOS assaults and conducted a comparison study between ANN, decision tree, antropy, and Bayesian. It is capable of generalizing from limited and imperfect data. Neural networks require more time to train and are not ideal for real-time detection.

3.3 Support Vector Machine (SVM)

Li et al. [1] suggested a customizable intelligent module for network intrusion prevention systems by integrating SNORT and firewall. By combining an SVM classifier with SNORT, the false alarm rate is minimized, boosting the accuracy of the intrusion prevention system. SVM produces outcomes that are simple to comprehend and efficient. The only issue is that SVM only handles binary classification. For binary classifiers, it fails to provide further information about the identified attack type.

3.4 Naïve Bayes

Alkasassbeh et al. [13] compiled a new dataset of DDOS assaults at various network levels. He employed three algorithms to identify DDoS attacks: Multilayer perceptron (MLP), naive Bayes, and random forest. MLP demonstrated the best accuracy (98.63%) when compared to other approaches. The advantage of naive Bayes is that it is simple to implement. On greater data points, naive Bayes might be used, but it adds complexity. The disadvantage is the requirement for probability data and the assumption of conditional independence.

3.5 Decision Tree

C4.5 decision tree algorithm is more stable than k-nearest neighbor algorithm, according to Ismanto and Wardoyo [19]. Experiment on tree intrusion detection is carried out, analyzing multi-layer perceptron (MLP), C4.5, and SVM classifiers, with C4.5 demonstrating that it is better in detection (99.05%) and has the shortest training time. The authors of [20] examined the C4.5, naive Bayes, and C5.0 in identifying DDOS assaults. When compared to the other two algorithms, C5.0 has the highest accuracy. Many studies have been conducted to conclude that C4.5 has been working well with increased accuracy, but C5.0 has begun to perform even better than C4.5 [20]. The decision tree has the advantage of increasing accuracy as the quantity of the datasets increases. The disadvantage of decision trees is that they are unstable when dealing with complex numerical datasets.

4 Conclusion

Following a comprehensive investigation, it is determined that network attacks are destructive and that proper intrusion detection systems must be implemented. As previously stated, each machine learning approach worked well with one or more features that might be tallied in accordance with the identification of an intrusion in a system. Some machine learning-based detection algorithms improve in accuracy as the number of datasets increases, while others have a lower false alarm rate, and yet others are good for port scans and probes. As a result, it can be stated that machine learning methods have both advantages and downsides in identifying DDOS attacks. Appropriate techniques should be chosen based on whether the DDOS attack is originating at the port, network, or application layer, as well as the type of datasets accessible.

References

1. Li H, Liu D (2010) Research on intelligent intrusion prevention system based on snort. In: International conference on computer, mechatronics, control and electronic engineering (CMCE), vol 1. IEEE, pp 251–253
2. Li J, Liu Y, Gu L (2010) DDos attack detection based on neural network. In: 2nd international symposium on aware computing (ISAC). IEEE, pp 196–199
3. Suresh M, Anitha R (2011) Evaluating machine learning algorithms for detecting DDoS attacks. Commun Comput Inform Sci 441–452. https://doi.org/10.1007/978-3-642-22540-6_42
4. Livadas C, Walsh R, Lapsley DE, Strayer WT (2006) Using machine learning techniques to identify botnet traffic. In: Proceedings of 2006 31st IEEE conference on local computer networks, pp 967–974
5. Suresh M, Anitha R (2011) Evaluating machine learning algorithms for detecting DDoS attacks. In: Wyld DC, Wozniak M, Chaki N, Meghanathan N, Nagamalai D (eds) Advances in network security and applications. CNSA 2011. Communications in computer and information science, vol 196. Springer, Berlin, Heidelberg
6. Wu Y-C, Tseng H-R, Yang W, Jan R-H (2011) DDoS detection and traceback with decision tree and grey relational analysis. Int J Ad Hoc Ubiquitous Comput 7(2)
7. Yusof AR, Udzir NI, Selamat A (2016) An evaluation on KNN-SVM algorithm for detection and prediction of DDoS attack. In: Fujita H, Ali M, Selamat A, Sasaki J, Kurematsu M (eds) Trends in applied knowledge-based systems and data science. IEA/AIE 2016. Lecture notes in computer science, vol 9799. Springer, Cham
8. Zekri M, El Kafhali S, Aboutabit N, Saadi Y (2017, October) DDoS attack detection using machine learning techniques in cloud computing environments. In: 2017 3rd international conference of cloud computing technologies and applications (CloudTech). IEEE, pp 1–7
9. Bujlow T, Riaz T, Pedersen JM (2012, January) A method for classification of network traffic based on C5.0 machine learning algorithm. In: 2012 international conference on computing, networking and communications (ICNC). IEEE, pp 237–241
10. Pandya R, Pandya J (2015) Article: C5.0 algorithm to improved decision tree with feature selection and reduced errorpruning. Int J Comput Appl 117(16):18–21
11. Bhuyan MH, Bhattacharyya DK, Kalita JK (2011) Surveying port scans and their detection methodologies. Comput J 54:1565–1581
12. Das V, Pathak V, Sharma S, Sreevathsan, Srikanth MVVNS, Gireesh Kumar T (2010) Network intrusion detection system based on machine learning algorithms. Int J Comput Sci Inform Technol (IJCSIT) 2(6)
13. Alkasassbeh M, Al-Naymat G, Hassanat ABA, Almseidin M (2016) Detecting distributed denial of service attacks using data mining techniques. Int J Adv Comput Sci Appl (IJACSA) 7(1)
14. Bains JK, Kaki KK, Sharma K (2013) Intrusion detection system with multi layer using Bayesian networks. Int J Comput Appl 67(5). ISSN 0975-8887
15. Gavrilis D, Dermatas E (2005) Real-time detection of distributed denial-of-service attacks using RBF networks and statistical features. Comput Netw 48:235–245. https://doi.org/10.1016/j.comnet.2004.08.014
16. Sofi I, Mahajan A, Mansotra V (2017) Machine learning techniques used for the detection and analysis of modern types of DDoS attacks. IRJET 4(6)
17. Balsrengadurali C, Saraswathi S (2013) Fuzzy based detection and prediction of DDoS attacks in IEEE 802.15.4 low rate wireless personal area network. IJCSI Int J Comput Sci 10(6)(1)
18. Bains JK, Kaki KK, Sharma K (2013) Intrusion detection system with multi-layer using Bayesian networks. Int J Comput Appl 67(5). ISSN 0975-8887
19. Ismanto H, Wardoyo R (2016) Comparison of running time between c4.5 and k-nearest neighbor (k-nn) algorithm on deciding mainstay area clustering. Int J Adv Intell Inform 2(1):1–6
20. Hariharan M, Abhishek HK, Prasad BG (2019) DDoS attack detection using C5.0 machine learning algorithm. Int J Wirel Microw Technol (IJWMT) 9(1):52–59

Review on 5G and Wi-Fi 6 Wireless Internet Connectivity

Satish Surendra Srivastava, Sanjay Makh, and P. R. Rodge

Abstract Upcoming wireless networks like 5G, Wi-Fi 6, and beyond are projected to be exceedingly complicated and dynamic. The rise of ultra-dense complex network deployment, high data rates, with new applications may necessitate a new wireless radio technology paradigm may offer several critical challenges for network administration, operations, and planning, including troubleshooting. Similarly, the generation as well as consumption of wireless data is shifting from persons to machine-oriented communications, considering future wireless network operations even more complex. Like a result, new approaches for deploying dispersed computation means with greater context awareness will become more important just for reduce the complication of future wireless network. Studies indicated that has been focused on wireless broadband connection of the 5th generation, called "5G," which is now implemented through mobile network operators. Unexpectedly, 'Wi-Fi 6', is the newest IEEE 802.1ax model inside group based on wireless local area network techniques including characteristics aimed for private, and edge-networks, has received far less attention. This paper examines the potential for cellular along with Wi-Fi networks for providing speedy wireless Internet connection.

Keywords 5G · 802.11 · Broadband · Internet · Wi-Fi 6

1 Introduction

We are witnessing the deployment with the descendants wireless communication technologies for mobile also Wi-Fi connectivity almost simultaneously. Although the 5th generation of cellular technology, so-called '5G,' has generated a lot of buzz across the world, so has the succeeding iteration based on the IEEE 802.11 wireless local access network (WLAN) standard, often known as 'Wi-Fi6', is a wireless

S. S. Srivastava (✉) · P. R. Rodge
Shivajirao S Jondhale College of Engineering, Dombivali East, Mumbai, India
e-mail: sksrivastava2370@gmail.com

S. Makh
Sant Gadagebaba University, Amaravati, Maharashtra, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_15

local access network (WLAN). Specified the productivity with societal aids, i.e., a successful transition provides, descendants wireless communication advancements remain required to extra move to a Digital Economy [1].

Wireless access in addition localization figured an undeniable part in permitting hundreds of billions of creative as well as prominent cyberspace applications that connect with the Internet along anybody, anyplace else within the earlier few years, as we witnessed the birth as to "information age" also 3rd engineering regime change. Such online application already has, that continues to have a major impact on how we live, do work, purchase, uses the media, provide instruction, transportation, health care, in addition communicate with the rest of the world. Wireless technology is being used by mobile phones, tablets, as well as laptops to enable unrestricted access to information, that is the most crucial component of how we work and live [2].

Even though there has been a heated discussion between technology and industry members in favor of cellular/Wi-Fi technology, perhaps the investigation community really hasn't properly considered the scope to which these technologies may interact as available choices in some connection to in addition also as more technologies that are related [3, 4].

Inside a 3G/4G world, most mobile customers choose to use data-intensive apps through Wi-Fi since it provides better performance as well as avoids the increased charges or data limitations involved in mobile services. Considering 5G's improved speed and the change to endless data plans for mobile, certain predict that end users would migrate overall data traffic over Wi-Fi toward cellular networks. Furthermore, the standards group 3GPP takes nowadays incorporated indoor broadband scenarios advocating and use of cellular technologies to facilitate connection in workplace environments, boosting the possibility of 5G directly competing with a crucial Wi-Fi scenario.

Inside this paper, we try comparing multiple viewpoints on the comparative advantages of 5G and Wi-Fi 6, for example, mentioned with in technical publications, as well as scientific articles. In Sect. 2, we begin by looking at the demand side trends that will distress wireless Internet access over the coming era. Sections 3 and 4 give a basic dependent review on 5G and Wi-Fi 6 for just a policy as well as economics listeners, followed by a making comparison of such technologies in Sect. 5. Section 6 concludes with some last observations.

2 Literature Review

Presently, mobile users are well-versed in cell phone (mobile) technologies. 5G technologies incorporate all forms of revolutionary structures, creating 5G mobile technology the utmost powerful also in high demand in the forthcoming era. 5G networks can be faster in the meantime they employ smaller frequencies (millimeter waves from 30 and 300 GHz). High-band 5G spectrum provides a predictable increase in speed, capacity, low latency, as well as quality [5].

5G is distinct owing to unique characteristics such as linking individuals and monitoring gadgets, objects, and machines. As a result, it is critical to understand where the company can capitalize on the advantages of 5G [6].

Upcoming mobile systems or next generation wireless networks are expected to provide high-speed access without regard to time or place. The incorporation of advanced technologies or new services by current network infrastructure is critical to the development of any technologies with significant data traffic with top quality of service of universal network basic structures of a system [7].

The advent of wireless communication's fifth generation (5G) may expand the amount of high frequency powered base stations as well as other devices. There is an urgent necessity for study on local heat developments on small surfaces, like skin or the eye, as well as at all environmental influence. According to quality analysis, upcoming studies' design and implementation must be greatly improved in order for them to be helpful for safety evaluation [8].

5G included a well-connected base network and radio access network (RAN). This core layer may potentially transition from fibers to mm-wave wireless communication, with linked base stations using high-capacity cable connections. Even as devices connected to the internet grows, a typical macro-cell might become overburdened by managing overheads in order to maintain communication through a large number of devices (near about 10 k/cell). As a result, the architecture needs to be simplified also modified in order to support greater signaling in addition payload expenses. P performance of such a future 5G architecture, using mm-wave RAN, has been documented in the Giga KOREA 5G development [9].

The subcarriers in the FBMC system are exposed to side-lobe suppression through running them through a filter bank, allowing them to outperform OFDM in terms of spectral efficiency. FBMC is not affected by inherent conflicts that may arise in between receiver and transmitter. The 5G NOW [10] group is now considering four waveforms, GFDM [11], UFMC [12], FBMC [13], and BFDM [14], aimed at 5G have to deliver an efficient air interface that is not reliant on strict orthogonality also synchronization constraints.

The IEEE 802.11ax standard provides the foundation for Wi-Fi 6. This technology will allow descendants Wi-Fi connection. Wi-Fi 6 will deliver the bandwidth, range, and speed required for successfully fulfill the growing and expanding demand for Wi-Fi infrastructure. Wi-Fi 6 is really the latest language used to describe Wi-Fi network devices including their connection [15].

For help clients identify equipment which deliver most recent Wi-Fi technology, the Wi-Fi association is introduced a simple vocabulary for naming distinct Wi-Fi generations [16]. Such terms may appear on devices as the device's name as well as summary. Wi-Fi devices that offer contemporary generation connection are developed from the 802.11ax standard also are referred to by means of Wi-Fi 6 devices [17].

A 802.11b was the first widely used Wi-Fi protocol. In place of new services such as video conferencing, smart homes, and augmented reality/virtual reality gain popularity, Wi-Fi networks get congested due to the growing number of smart devices. As a result, enhancing spectrum efficiency seems to be more important over raising

theoretical rate for the next Wi-Fi version, 802.11ax, it addresses attempts to enable access to more Stations (STAs) [18].

W-Fi 6 will make it easier for customers to understand and use high-performance Wi-Fi technology. Further crucially, Wi-Fi 6 raises network efficacy toward a level that will allow for contemporary novel applications in practically every part. Only a rare of these applications are previously visible [19].

The analysis' main purpose is to show how 5G also Wi-Fi 6 could alter the aggressive dynamics among the cellular with Wi-Fi technology groups. In the direction of decide if one technology appears to be the dominating choice otherwise whether both technologies will continue to be essential also frequently complimentary instruments aimed at satisfying wireless demands in the coming years.

3 Technical Aspects of 5G

The 3GPP is an industry group that develops worldwide specifications for 3G, 4G, and 5G technology, is in charge of cellular standardization efforts. Aimed at 5G, the initial requirement produced by 3GPP because phase 1 (Release 15) identifies three important technological use cases (3GPP, Year-2019).

D2D connection is required in a cellular system to boost cell capacity also provide different proximity services (ProSe) (example, retail/restaurant, bus otherwise cab availability, ads in a region, browsing social networks within a huge crowd.). During its release-12, 3GPP standardized LTE-Direct (LTE-D) [20], and mm-wave D2D communication is required for 5G. Mm-wave D2D communication is demonstrated in [21] mobile devices equipped by electrically directed antennas also beamforming technologies. Concurrent transmission between D2D users as well as the base station will not cause much interference due to highly directional antennas and beam-forming technologies.

Although bad network planning typically results in major cost implications [22–25], the large deployment of tiny cells provides a new set of issues relating to spectrum management, energy efficiency, and the logistics of installing backhaul. The disaggregation of tiny cell structures, based on the virtualization concepts outlined far along in reference to Fig. 1, is one approach to addressing some of these challenges.

Its use of directional antennas in 5G is a vital feature that allows a single stream of information to be targeted toward the user instead of being delivered widely, considerably decreasing congestion at other cells. By triangulating the position of the user device in conjunction with mMIMO, a BS may then predict the most effective path to convey information packets based on decreased interference. A recent study discovered a viable 5G business case for eMBB from 2020 to 2030 [27].

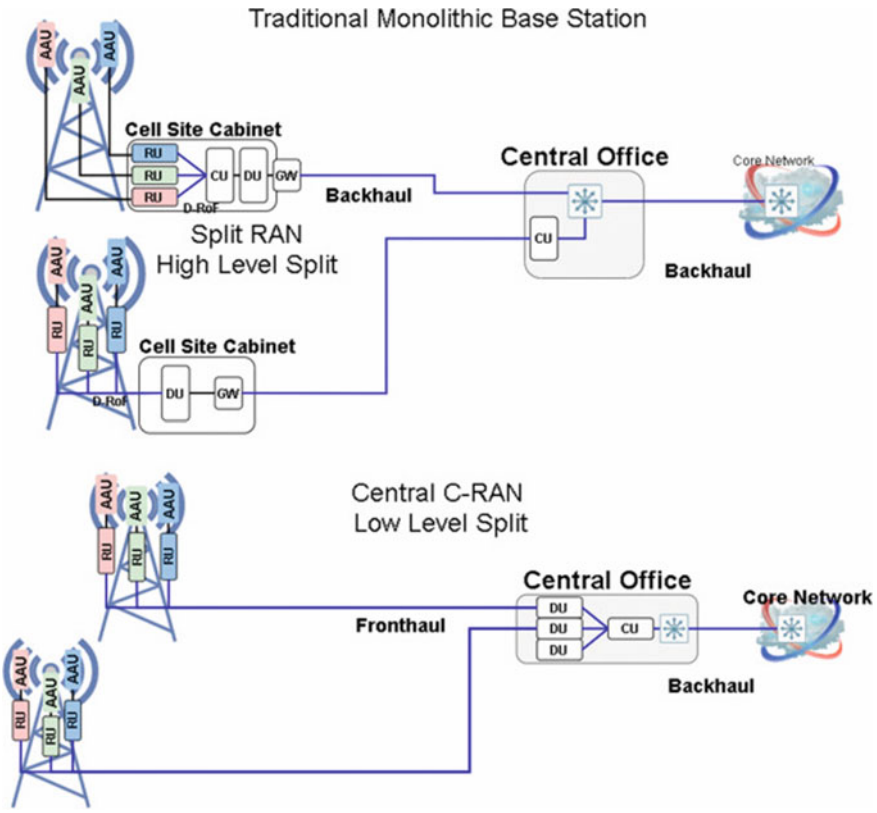


Fig. 1 Cellular RAN configuration evolution [26]

4 Technical Aspects of Wi-Fi 6

Throughout this section, we will discuss the key aspects of Wi-Fi 6 technology. Wi-Fi development has been led by two distinct industry organizations. To begin, IEEE Project 802 is in the process of being developed group in charge of various networking standards, with the Wi-Fi technology suite. Another, the Wi-Fi Association is a profitless organization comprise an international network of corporations assigned by assuring inter-operability as well as promoting and certifying various Wi-Fi technologies. This requires in charge of in cooperation technical factors, for example developing extra standards aimed at products, i.e., Wi-Fi mesh networks, also administration problems, like working with policymakers about appropriate spectrum allotments.

IEEE 802.11ax, sometimes referred by means of 'Wi-Fi 6', seems to be the initial addition towards the Wi-Fi family intended move out tiny interior contexts, with the goal of optimizing efficiency in large outdoor installations. Even though it advances

minimal data rate about 37% on Wi-Fi 5, intends to deliver a 4× gain in throughput with spectrum efficiency for installations that is dense with additional features.

Wi-Fi mesh systems are becoming increasingly common, where instead of a single AP, the system consists of a primary hub with several connected nodes geographically spread around a building or residence that are capable of gathering as well as rebroadcasting information [28]. This method aids in the elimination of locations having low signal coverage, boosting overall speed as dependability for consumers. With ease which Wi-Fi can be deployed contrasts sharply with the technological constraints of 5G deployment [29].

5 5G versus Wi-Fi 6—A Comparison as Well as Contrast

Throughout this study, differentiating 5G and Wi-Fi 6 in terms of technological assets, spectrum use, costing as well as business model, also easiness of installation then necessary expertise level. Such comparison categories were chosen because account in both engineering in addition economic position when it comes to the technological innovation. 5G and Wi-Fi 6 provide major performance improvements over previous technological generations, with significantly increased connection speed, higher device numbers, and lower latency.

In this section, sort out some of these competing factors in order to gain a better understanding of by what means 5G and Wi-Fi 6 can affect the nature of competition of cellular with Wi-Fi-based technology. Table 1 gives higher summarized comparison between 5G with Wi-Fi 6 across a variety of technical and economic parameters.

6 Conclusion

We addressed the wireless Internet connectivity discussion in this paper by presenting a novel review in between two primary methods used in supply of future era wireless broadband: 5G, Wi-Fi 6. The investigation emphasizes in what way the upcoming of 5G, and Wi-Fi 6 must recognize in connection with how old lines of cellular besides Wi-Fi technologies affected the growth for wireless networking then what it really possibly will signify aimed at the tomorrow.

The analysis' main purpose is to show in what way 5G besides Wi-Fi 6 could modify the changing aspects of competition among cellular as well as Wi-Fi technology generations. To decide if one technology appears to be the dominating selection otherwise regardless in cooperation technologies will continue to be essential and frequently complimentary instruments for satisfying wireless demands in the upcoming era. The technological features, both new cellular and Wi-Fi generations aspire towards deliver higher spectrally effective radio interfaces provide an improved user experience. However, find that, like earlier cellular generations, 5G is

Table 1 Comparisons of 5G and Wi-Fi 6 characteristics

Variable	5G	Wi-Fi 6
Peak data rate	1 Gbps (UL), 2 Gbps (DL)	5 Gbps (UL), 10 Gbps 8 × 8 (DL)
MU-MIMO	128 × 128	8 × 8
Coverage range	Small cells have a range of 100–300 m, while macro cells have a range of tens of kilometers	less than 50 m inside, 300 m outdoors
Carrier aggregation	Yes	Yes 40, 80, 160 (or 80 + 80)
Inter-cell interference	Controlled	Mainly uncontrolled
Channel access scheme	OFDMA	OFDMA
License type	Mostly licensed	Unlicensed
General bands	Low, mid and high	Low and mid
Specific frequencies	High-band (~ 24–29 GHz) mid band (1–7 GHz), low-band (< 1 GHz)	60, 6, 5, and 2.4 GHz
Channel bandwidth	20, 40, 80 and 100 MHz	20, 40, 80 and 160 MHz
Revenue model	Data service billing might be pre- or post-paid	Perhaps a service, a "free" facility, or a pure WLAN with no outside connection
User equipment price	High	Low
Public versus private	Conventionally publicly provided through an MNO	Traditionally privately provided
Chip/modem cost	High	Low
Data cost	Monthly subscription ($5–20)	Free
Deployment approach	Controlled and managed	Uncontrolled also mostly unmanaged
Installation skill level	High	Low
Development skill level	High	Low

still focused on providing consumers with great mobility. While Wi-Fi continues to be geared at offering nomadic high-capacity hotspots that are quickly deployable.

References

1. Bauer JM (2018) The Internet and income inequality: socio-economic challenges in a hyperconnected society. Telecommun Policy SI Interconnecting 42:333–343. https://doi.org/10.1016/j.telpol.2017.05.009
2. Pahlavan K, Krishnamurthy P (2021) Evolution and impact of Wi-Fi technology and applications: a historical perspective. Int J Wirel Inf Netw 28:3–19
3. Light R (2019) Will 5G kill WiFi? 3.2.22. https://www.lightreading.com/mobile/5g/will-5g-kill-wifi-qualcomm-thinks-it-just-might/d/d-id/749618. Last accessed on 2022/03/02

4. Bloomberg (2017) A world without Wi-Fi looks possible as unlimited plans rise. Bloomberg.com
5. Pujari V, Patil R, Tambe K (2021) Research paper on future of 5G wireless system
6. Dangi R, Lalwani P, Choudhary G, You I, Pau G (2022) Study and investigation on 5G technology: a systematic review. Sensors 22:26. https://doi.org/10.3390/s22010026
7. Le NT, Hossain MA, Islam A, Kim D, Choi Y-J, Jang YM (2016) Survey of promising technologies for 5G networks. Mob Inform Syst 2016:1–25. Hindawi Publishing Corporation
8. Simkó M, Mattsson MO (2019) 5G wireless communication and health effects—a pragmatic review based on available studies regarding 6 to 100 GHz. Int J Environ Res Public Health 16(18):3406
9. Bae JS et al (2014) Architecture and performance evaluation of MmWave based 5G mobile communication system. In: IEEE ICTC 2014, pp 847–851
10. 5G now deliverables. Accessed on 2022/07/03. http://www.5gnow.eu/wp-content/uploads/2015/04/5GNOW_D3.2_v1.3.pdf
11. Fettweis G, Krondorf M, Bittner S (2009) GFDM—generalized frequency division multiplexing. In: 69th vehicular technology conference. IEEE, pp 1–4
12. Mukherjee M et al (2015) Reduced out-of-band radiation-based filter optimization for UFMC systems in 5G. In: Wireless communications and mobile computing conference, IWCMC, pp 1150–1155
13. Van der Neut N et al (2014) PAPR reduction in FBMC systems using a smart gradient-project active constellation extension method. In: 21st international conference on telecommunications, ICT, pp 134–139
14. Wunder G, Gorgani SA, Ahmed SS (2015) Waveform optimization using trapezoidal pulses for 5G random access with short message support. In: IEEE 16th international workshop on signal processing advances in wireless communications, pp 76–80
15. Krishan R (2019) Wi-Fi 6 technology—a review. Int J Modern Electron Commun Eng (IJMECE) 7(4):11–13
16. Wi-Fi 6: High performance, next generation Wi-Fi, White Paper, Wi-Fi Alliance (2018)
17. IEEE 802.11ax: The Sixth Generation of Wi-Fi, Technical White Paper, Cisco Public (2018)
18. Wi-Fi 6 Technology and Evolution White Paper, ZTE (2020)
19. Shaji G, Hovan G (2020) A review of Wi-Fi 6, the revolution of 6th generation Wi-Fi technology. Int J Eng Sci 10(9):56–65
20. LTE Direct Workshop White Paper-Qualcomm Inc
21. Qiao J et al (2015) Enabling device-to-device communications in millimeter-wave 5G cellular networks. IEEE Commun Mag 209–215
22. Haddaji N, Bayati A, Nguyen K, Cheriet M (2018) BackHauling-as-a-Service (BHaaS) for 5G optical sliced networks: an optimized TCO approach. J Lightwave Technol 36:4006–4017
23. Taufique A, Jaber M, Imran A, Dawy Z, Yacoub E (2017) Planning wireless cellular networks of future: outlook, challenges and opportunities. IEEE Access 5:4821–4845
24. Wisely D, Wang N, Tafazolli R (2018) Capacity and costs for 5G networks in dense urban areas. IET Commun 12:2502–2510
25. Yaghoubi F, Mahloo M, Wosinska L, Monti P, de Souza Farias F, Costa JCWA, Chen J (2018) A techno-economic framework for 5G transport networks. IEEE Wirel Commun 25:56–63
26. Alsharif MH, Nordin R (2017) Evolution towards fifth generation (5G) wireless networks: current trends and challenges in the deployment of millimetre wave, massive MIMO, and small cells. Telecommun Syst 64:617–637
27. Rendon Schneir J, Ajibulu A, Konstantinou K, Bradford J, Zimmermann G, Droste H, Canto R (2019) A business case for 5G mobile broadband in a dense urban area. Telecommun Policy
28. Navío-Marco J, Arévalo-Aguirre A, Pérez-Leal R (2019) WiFi4EU: techno-economic analysis of a key European Commission initiative for public connectivity. Telecommun Policy 43:520–530
29. Forge S, Vu K (2020) Forming a 5G strategy for developing countries: a note for policy makers. Telecommun Policy 44:101975

Handwriting Recognition and Conversion Using Neural Networks

Aditya Saini, Kunal Sant, Sumeet Swain, and Neha Deshmukh

Abstract In the current world of automation, everything is getting atomized as reducing manual labor is the key to efficiency. Here, we focus on offline handwriting recognition. Handwriting recognition is the process of extracting text from handwritten scripts, this is also known as offline handwriting recognition. The purpose here is to attempt to improve the accuracy and efficiency of the system using neural networks and datasets used for training the model, as well as detecting and identifying the characters and exporting them in a text format. The proposed system can be used to recognize handwritten characters and convert them into the text from the scanned image of a page.

Keywords Neural networks · Handwriting recognition · Long short-term memory · Optical character recognition · Convolutional neural network

1 Introduction

Handwriting recognition is a subject which existed for a long period of time, but the accuracy of the process can be improved by the usage of more refined technology and better techniques that reduce the error rates and the difference between the actual text and the detected text of the system [1]. Handwriting recognition is the process of extracting text from handwritten scripts, and this is also known as offline handwriting recognition [2]. In offline handwriting recognition, the image of a script is captured using a scanner and sent forward for further processing [3].

A. Saini (✉) · K. Sant · S. Swain · N. Deshmukh
A P Shah Institute of Technology, Thane(w), Maharashtra, India
e-mail: adityasaini@apsit.edu.in

K. Sant
e-mail: kunalsant@apsit.edu.in

S. Swain
e-mail: sumeetswain@apsit.edu.in

N. Deshmukh
e-mail: npdeshmukh@apsit.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_16

This is a step into the world of automation and can help us improve efficiency in this field. Handwriting recognition works on similar techniques of text detection and recognition. Text detection includes various techniques to extract text from an image, such as the texture detection method and connected component method. In texture detection, the text is extracted, which is detected by texture difference from the background, and text is registered by further processing of the information. In the connected component method, the text is extracted using the relation between pixel connectivity and pixel intensity to detect text patterns, and checks are made to clear non-character elements [4]. The next step in extracting text from an image is text recognition which is done by optical character recognition (OCR) which depends on identifying the character from the detected text, which can be done by training the model using a dataset that gives the model the basic characteristics of characters like the height, width, shape and size and font style. This includes all the alphanumeric characters, 10 digits (0–9), and 26 alphabets with uppercase and lowercase for a total of 62 characters [3]. Handwritten characters are more challenging to recognize as compared to printed characters as they carry more parameters like the standard and cursive style of writing [4]. These challenges can be overcome by using new technologies like long short-term memory (LSTM), recurrent neural network (RNN), convolutional neural network (CNN), and different training methods with various datasets [5]. The handwriting can be extracted from an image through this set of processes.

RNNs are popularly used in generating data for sequences in music text and motion capture, and they can be trained by real data. A particular property of RNNs is that they can create new data with the help of existing data, i.e., interpolation (a statistical method).

In principle, a sufficiently large RNN should be sufficient to generate strings of arbitrary complexity. In practice, however, it has been observed that standard RNNs cannot store information about past inputs for very long periods of time due to their short memory. This has a significant effect on their ability to perform well in long-span structures. This is known as “amnesia” behavior, which makes them susceptible to instability when creating strings. The suggested remedy is to introduce noise into the predictions before feeding them back to the model, which will increase the internal efficiency of the model to handle any unexpected inputs that may arise. However, we believe that increasing system memory is a deeper and more efficient solution.

The long-term short-term memory network is LSTM. A long-term memory network is a type of cyclic neural network (RNN). The LSTM architecture is driven by error stream analysis in existing replay neural networks. Memory blocks are a set of recursively connected blocks used in the LSTM layer. These blocks can be thought of as distinguishable versions of memory chips in digital computers. Each of these blocks contains one or more repeatedly connected memory cells and three multipliers: input, output, and forget gate. They provide persistent analogies of cell write, read, and reset operations. The interaction between the network and the cell can only happen through ports (Fig. 1).

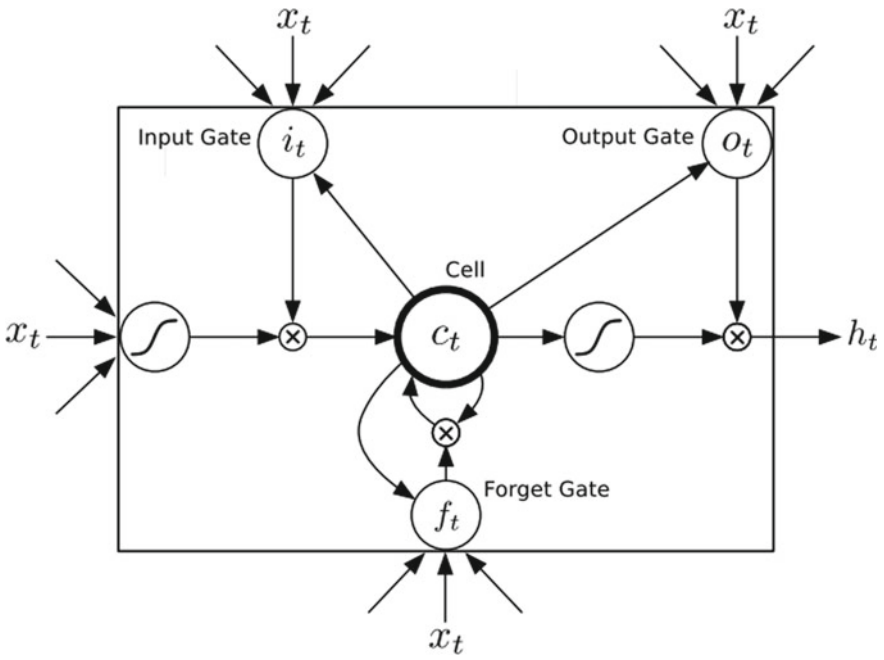


Fig. 1 Long short-term memory cell, along with its three multiplicative units, that is: the input, the output and the forget gates

2 Literature Review

In paper [3], use of techniques such as preprocessing, segmentation, and feature extraction that are used to improve the accuracy of text recognition and identification. In paper [4], the work of different text detection methods on how they extract text from an image such as MSER and canny edge integration, stroke width variation, and OCR for text recognition. In paper [5], combination of CNN and LSTM showing a better performance for offline handwriting recognition. In paper [6], how using CNNs and their training with GPUs to train models can reduce character recognition error rates in offline handwriting. In paper [7], we learned that LSTM-RNN with backpropagation through time training can be used to speed up the process while increasing the accuracy of the recognition process.

3 Limitations

Quality factors such as noise in the image, blurred images, and in general low quality of the image and text can hinder the performance, leading to mistakes in the end result [4].

Similar looking characters can lead to misclassification by the system (1, l, i, I) (o, 0, O) (2, z, Z) (s, S, 5) (A, 4) [6]. Currently, the system works on an image of a set resolution and aspect ratio for the working. The system can only recognize characters that are trained to the system, so if the system is trained for alphabets, then it won't be able to recognize numbers and special characters and punctuation marks.

4 Previous Work

Researchers have tried increasing the efficiency of the text detection process by increasing the size of the sliding window, which is used to detect text [4]. Neural networks with various training techniques with datasets have also improved the speed and accuracy of the process [7].

5 Implementation

Handwriting to text

CNN: The user's data is initially placed into CNN layers. To extract the features from the image, CNN layers are employed. Five input layers are used, with the first two applying a 5×5 filter and the last three applying a 3×3 filter for convolution operations. A non-linear RELU function is implemented after the input passes through the five layers of CNN. Finally, image sections are summarized using a layer, and a reduced output is generated. Despite the fact that the image is shrunk by two in each layer, a 32×256 feature map is added.

RNN: The RNN passes important information to the feature sequence, which comprises 256 features each time step. The popular long short-term memory (LSTM) implementation of RNNs is employed because it can transmit information over longer distances and has more powerful training properties than a standard RNN. The RNN output sequence is mapped into a 32×38 matrix. The IAM dataset contains a total of 79 different characters, plus one extra character for the CTC process, for a total of 80 entries for each of the 32 time steps.

Connectionist Temporal Classification (CTC): The CTC is supplied to the RNN output matrix and the ground truth text during training the neural network, which computes the loss value. The CTC is only provided the matrix while inferring, and it decodes the final text using that information. Ground truth and recognized text can only be 32 characters long.

Input Data: The input data is a 128×32 pixel picture. Because the images in the collection aren't exactly this size, we resize them to a width of 128 pixels or a height of 32 pixels. The input is then transformed into a 128×32 target image (white).

Data augmentation can be implemented by arbitrarily moving or enlarging the image rather than aligning it to the left.

CNN Output: Each entry in the output has 256 characteristics. The RNN layers further process the features that we have received. However, some of the features already have a strong relationship with the input image's high-level qualities. Some features have a strong relationship with duplicate characters (for example, "tt"), specific characters (for example, "e"), or character qualities such as loops and curves (for example, in handwritten "l"s or "e"s).

Implementation using Tensorflow:

There are four modules in the implementation.

To prepare the photos from the dataset for the neural network, use SamplePreprocessor.py.

DataLoader.py is a Python module that reads samples, organizes them into batches, and provides an iterator interface for traversing the data.

Model.py: Creates a model from a given architecture, loads and saves models, controls the session at the same time, and provides a training and inference interface.

main.py: This file combines all of the previously described modules. The other source files are dealing with basic file IO, input and output (DataLoader.py), and picture processing, so we simply look at Model.py (SamplePreprocessor.py) (Figs. 2 and 3).

Text to Handwriting

Input Data: This model will be fed with computer-generated text as its input. We used online handwritten data to assess the model's ability to construct plausible real-valued sequences, as opposed to offline handwriting, which only has page pictures. We got all of our information from IAM-OnDB, which is an online handwriting database.

LSTM: Backpropagation via time was used by the LSTM method to build a full gradient. When using the whole gradient to train LSTM, one issue is that the derivatives might get too big, causing numerical issues. To avoid this, we clipped the

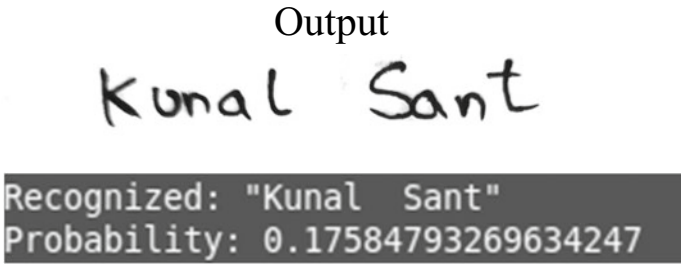


Fig. 2 The input text "Kunal Sant" given to the model and its output, along with the accuracy that we have achieved, is shown

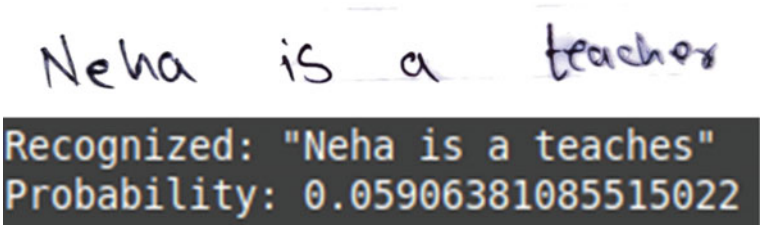


Fig. 3 The input text "Neha is a teacher" given to the model and its output, along with the accuracy that we have achieved, is shown

derivative of the loss with respect to the network inputs to the LSTM layers in all of the tests such that it stayed within a preset range.

6 Conclusion

We want to improve the efficiency of handwriting recognition. We talked about a neural network that can recognize text in photos. The neural network outputs a character-probability matrix and has five CNN and two RNN layers. This matrix is either used to calculate CTC loss or to decode CTC. TensorFlow is used in the implementation.

Acknowledgements We thank Prof. Neha Deshmukh for helping us in creating this research paper and guiding us throughout this project. We also thank Prof. Kiran Deshpande, Head of the Department of Information Technology, for his help in our work.

References

1. Chherawala Y, Roy PP, Cheriet M (2015) Feature set evaluation for offline handwriting recognition systems: application to the recurrent neural network model. IEEE Trans Cybern 46(12)
2. Senior A, Robinson T (1998) An off-line cursive handwriting recognition system. IEEE Trans Pattern Anal Mach Intell 20(3)
3. Pradee J, Srinivasan E, Himavathi S (2011) Diagonal based feature extraction for handwritten character recognition system using neural network. In: 2011 3rd international conference on electronics computer technology (ICECT), vol 4
4. Islam MR, Mondal C, Azam MK, Islam ASMJ (2016) Text detection and recognition using enhanced MSER detection and a novel OCR technique. In: 2016 5th international conference on informatics electronics and vision (ICIEV)
5. Suryani D, Doetsch P, Ney H (2016) On the benefits of convolutional neural network combinations in offline handwriting recognition. In: 2016 15th international conference on frontiers in handwriting recognition (ICFHR)

6. Ciresan DC, Meier U, Gambardella LM, Schmidhuber J (2011) Convolutional neural network committees for hand-written character classification. In: International conference on document analysis and recognition
7. Doetsch P, Kozielski M, Ney H (2014) Fast and robust training of recurrent neural networks for offline handwriting recognition. 2014 14th international conference on frontiers in handwriting recognition (ICFHR)
8. Hu J, Brown MK, Turin W (1996) HMM based online handwriting recognition. IEEE Trans Pattern Anal Mach Intell 18(10)
9. Graves A (2013) Generating sequences with recurrent neural networks, pp 1–43. arXiv preprint arXiv:1308.0850. Retrieved from https://arxiv.org/abs/1308.0850
10. Tappert CC, Suen CY, Wakahara T (1990) The state of the art in online handwriting recognition. IEEE Trans Pattern Anal Mach Intell 12(8)
11. Zheng Y, Li H, Doermann D (2004) Machine printed text and handwriting identification in noisy document images. Pattern Anal Mach Intell 26(3)
12. Graves A, Liwicki M, Fernández S, Bertolami R, Bunke H, Schmidhuber J (2009) A novel connectionist system for unconstrained handwriting recognition. IEEE Trans Pattern Anal Mach Intell 31(5)
13. Bahlmann C, Haasdonk B, Burkhardt H (2002) Online handwriting recognition with support vector machines—a kernel approach. In: Proceedings of the 8th international workshop on frontiers in handwriting recognition (IWFHR)
14. Marti U, Bunke H (1999) A full English sentence database for off-line handwriting recognition. In: Proceedings of 5th international conference on document analysis and recognition, Bangalore, India
15. Xu L, Krzyzak A, Suen CY (1992) Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans Syst Man Cybern 22

On Cordial Totally Magic Labeling of Flower Graphs

R. Parameswari and C. Jayalakshmi

Abstract A graph $G(V, E)$ is called total magic cordial graph if it owns up the labeling called total magic cordial labeling which is defined as, a function $f : V(G) \cup E(G) \to \{0, 1\}$ such that $f(u) + f(v) + f(uv) \equiv C \pmod 2$ for all arcs $uv \in E(G)$ provided the condition $|f_0 - f_1| \leq 1$ is hold, where $f_0 = u_0 + e_0$ and $f_1 = u_1 + e_1$ where $u_i, e_i, i \in \{0, 1\}$ denote the nodes and arcs, respectively. This research article found that octopus graph, vanessa graph, lilly graph, and lotus graph admits the above mentioned labeling for all values of n.

Keywords Graph labeling · Total magic cordial labeling, · Octopus graph · Vanessa graph · Lilly graph and Lotus graph

AMS Classification 05C78

1 Introduction

This article deals with simple and undirected graphs. Harary [1] was followed for standard notations and terminology, and Gallian [2] dynamic survey is the best to refer about graph labeling. Cahit [3] 1987, prefaced the conception of cordial labeling. Totally magic cordial labeling was introduced by Cahit [4]. Cahit verified that the subsequent graphs have a TMC labeling: $K_{m,n}(m, n > 1)$, trees, complete graph if and only if $n = 2, 3, 5$ or 6. Total magic cordial labeling is also extended for directed graphs. Parameswari and Rajeswari [5] have proved that strong regular graph called Paley digraphs admits the same. They also [6] showed that (P_n, K_n), Gear and shadow graphs are total magic cordial. Parameswari [7] exhibited the same for square and shadow of bistar graph. Jeyanthi et al. [8] proved that flower graph Fl_n for $(n \geq 3)$, ladder graph, and graph with $|p - q| \leq 1$ own up total magic cordial labeling. In

R. Parameswari (✉) · C. Jayalakshmi
Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
e-mail: paramsumesh2011@gmail.com

C. Jayalakshmi
e-mail: jayachandru003@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_17

this article, we found total magic cordial labeling of octopus graph, vanessa graph, lilly graph, tortoise graph, and lotus graph.

1.1 Preliminaries

Graph labeling

Labeling is transferring numerals to the nodes or arcs, or both, with assured conditions.

Total magic cordial labeling

A function $f : V(G) \cup E(G) \to \{0, 1\}$ such that $f(a) + f(b) + f(ab) \equiv C(\text{mod}2)$ for all arcs $ab \in E(G)$ endowed with the condition $|f_0 - f_1| \leq 1$ is hold, where $f_0 = u_0 + e_0$ and $f_1 = u_1 + e_1$ is affirmed total magic cordial labeling.

Path graph

A path of a graph G is a finite order of arcs which joins a sequence of nodes which are all distinct. The path graph P_n has n nodes and $n - 1$ arcs.

Fan graph

Fan graph F_n, $(n \geq 2)$ is obtained by union of all nodes of the path P_n to a supplementary vertex (center) contains $n + 1$ nodes and $2n - 1$ arcs.

Star graph

Star $K_{1,n}$ is a complete bipartite graph. It has $n + 1$ *nodes and n arcs.*

Octopus graph

The octopus graph O_n, $(n \geq 2)$ is constructed by a fan graph F_n, $(n \geq 2)$ joining a star graph $K_{1,n}$ with sharing a common vertex which is the center of the star. In other words, $O_n = F_n + K_{1,n}$, n is a positive integer. It has $2n + 1$ nodes, $3n - 1$ arcs.

Vanessa graph

The vanessa graph V_n, $(n \geq 2)$ is constructed by two fan graphs F_n, $(n \geq 2)$ of the same order joining a star graph $K_{1,n}$ with sharing a common vertex which is the center of the star. In other words, $V_n = 2F_n + K_{1,n}$, where n is any positive integer. It has $3n + 1$ nodes and $5n - 2$ arcs.

Lilly graph

The lilly graph I_n, $(n \geq 2)$ is obtained by joining two star graph $2K_{1,n}$, $(n \geq 2)$ and two path graph $2P_n$, $(n \geq 2)$ with sharing a common node. In other words, $I_n = 2K_{1,n} + 2P_n$, where n is any positive integer. It has $4n - 1$ nodes and $4n - 2$ arcs.

Lotus graph

In a shell graph add a node in between each pair of adjacent nodes on the cycle and adding an arc in apex and two or more chords is known as lotus graph Lo_n. It has $4n + 9$ nodes and $6n + 11$ arcs.

2 Main Results

Theorem 2.1 *Let $G = O_n$ be a graph constructed by a fan graph joining a star graph. O_n has 2n + 1 nodes and 3n – 1 arcs. If f: $V \cup E \rightarrow$ {0, 1} be a function such that f(u) + f(v) + f(uv) = C(mod 2) for every uv ∈ E(G) then G is total magic cordial graph.*

Proof $G = O_n$ ($n \geq 2$) is an octopus graph, where $|V| = 2n + 1$ nodes and $|E| = 3n - 1$ arcs. The node set is affirmed as $V = \cup_{i=1}^{3} V_i$ where $V_1 = \{u_0\}$,$V_2 = \{u_i \ 1 \leq i \leq n\}$ and $V_3 = \{v_i \ 1 \leq i \leq n\}$.The edge set is $E = \cup_{i=1}^{3} E_i$, where $E_1 = \{u_0u_i \ 1 \leq i \leq n\}$, $E_2 = \{u_0v_i \ 1 \leq i \leq n\}$ and $E_3 = \{v_iv_{i+1} \ 1 \leq i \leq n-1\}$.

A function $f : V \cup E \rightarrow \{0, 1\}$ is as follows:

$$f(u_0) = 1 \tag{1}$$

$$f(u_i) = \begin{cases} 1 \ i \text{ is odd} \\ 0 \text{ i is even} \end{cases} \tag{2}$$

$$f(v_i) = \begin{cases} 0 \ i \text{ is odd} \\ 1 \ i \text{ is even} \end{cases} \tag{3}$$

$$f(u_0u_i) = 0 \quad \text{for all } i \tag{4}$$

$$f(u_0v_i) = \begin{cases} 0 \ i \text{ is odd} \\ 1 \ i \text{ is even} \end{cases} \tag{5}$$

$$f(v_iv_{i+1}) = 1 \quad \text{for } 1 \leq i \leq n-1 \tag{6}$$

To check that $|f_0 - f_1| \leq 1$.That is, $f_0 = u_0 + e_0$ and $f_1 = u_1 + e_1$ where u_i,e_i, $i \in \{0, 1\}$ are total count of nodes and arcs labeled as i, respectively (Fig. 1).

Case (i): n is odd

$$f_0 = \left(\frac{n-1}{2}\right) + \left(\frac{n+1}{2}\right) + n + \left(\frac{n+1}{2}\right) = \left(\frac{5n+1}{2}\right)$$

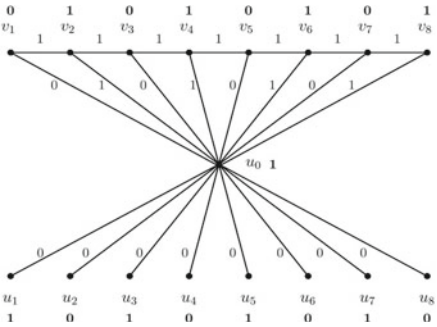


Fig. 1 Illustration of TMC labeling of O_8

$$f_1 = 1 + \left(\frac{n+1}{2}\right) + \left(\frac{n-1}{2}\right) + \left(\frac{n-1}{2}\right) + (n-1)$$
$$= \left(\frac{5n-1}{2}\right)$$
$$\therefore |f_0 - f_1| = \left|\left(\frac{5n+1}{2}\right) - \left(\frac{5n-1}{2}\right)\right| = 1$$

Case (ii): n is even

$$f_0 = \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + n + \left(\frac{n}{2}\right) = \left(\frac{5n}{2}\right)$$
$$f_1 = 1 + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + (n-1) = \left(\frac{5n}{2}\right)$$
$$\therefore |f_0 - f_1| = \left|\left(\frac{5n}{2}\right) - \left(\frac{5n}{2}\right)\right| = 0$$

In both cases, it is verified that $|f_0 - f_1| \leq 1$.
Hence, octopus graph $O_n (n \geq 2)$ admits the labeling for every n.

Theorem 2.2 *Let $G = V_n$ be a graph constructed by two fan graphs of the same order joining a star graph with sharing a common vertex which is the center of the star. V_n has $3n + 1$ nodes and $5n - 2$ arcs. If we define $f: V \cup E \to \{0, 1\}$ be a function such that $f(u) + f(v) + f(uv) = C (\text{mod } 2)$ for every $uv \in E(G)$ then G is total magic cordial graph.*

Proof $G = V_n$ $(n \geq 2)$ is a Vanessa graph with $|V| = 3n + 1$ nodes and $|E| = 5n - 2$ arcs, where V is denoted as $V = \cup_{i=1}^{4} V_i$, $V_1 = \{u_0\}$,$V_2 = \{u_i \text{ for } 1 \leq i \leq n\}$, $V_3 = \{v_i \text{ for } 1 \leq i \leq n\}$ and $V_4 = \left\{v_i' \text{ for } 1 \leq i \leq n\right\}$, and E is denoted as $E = \cup_{i=1}^{5} E_i$ where $E_1 = \{u_0 u_i \text{ for } 1 \leq i \leq n\}$, $E_2 = \{u_0 v_i \text{ for } 1 \leq i \leq n\}$, $E_3 = \left\{u_0 v_i' \text{ for } 1 \leq i \leq n\right\}$, $E_4 = \{v_i v_{i+1} \text{ for } 1 \leq i \leq n-1\}$ and $E_5 = \left\{v_i' v_{i+1}' \text{ for } 1 \leq i \leq n-1\right\}$.

Let $f : V \cup E \to \{0, 1\}$ as follows:

$$f(u_0) = 1 \tag{7}$$

$$f(u_i) = \begin{cases} 0 & i \text{ is odd} \\ 1 & i \text{ is even} \end{cases} \tag{8}$$

$$f(v_i) = \begin{cases} 0 & i \text{ is odd} \\ 1 & i \text{ is even} \end{cases} \tag{9}$$

$$f\left(v_i'\right) = \begin{cases} 1 & i \text{ is odd} \\ 0 & i \text{ is even} \end{cases} \tag{10}$$

$$f(u_0u_i) = \begin{cases} 0 & i \text{ is odd} \\ 1 & i \text{ is even} \end{cases} \tag{11}$$

$$f(u_0v_i) = 0 \quad \text{for } 1 \le i \le n \tag{12}$$

$$f\left(u_0v_i'\right) = 1 \quad \text{for } 1 \le i \le n \tag{13}$$

$$f(v_iv_{i+1}) = 0 \quad \text{for } 1 \le i \le n-1 \tag{14}$$

$$f\left(v_i'v_{i+1}'\right) = 1 \quad \text{for } 1 \le i \le n-1 \tag{15}$$

To prove that $G = V_n (n \ge 2)$ owns up total magic cordial labeling, we require that sum of nodes and arcs labeled 0 and sum of nodes and arcs labeled 1 differ at most by 1. That is, to prove that $|f_0 - f_1| \le 1$ (Fig. 2).

Case (i): n is odd

$$f_0 = \left(\frac{n+1}{2}\right) + \left(\frac{n+1}{2}\right) + \left(\frac{n-1}{2}\right) + \left(\frac{n+1}{2}\right) + n + (n-1) = 4n$$

$$f_1 = 1 + \left(\frac{n-1}{2}\right) + \left(\frac{n-1}{2}\right) + \left(\frac{n+1}{2}\right) + \left(\frac{n-1}{2}\right) + n + (n-1) = 4n - 1$$

$$\therefore |f_0 - f_1| = |4n - (4n-1)| = 1$$

Case (ii): n is even

$$f_0 = \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + n + (n-1) = 4n - 1$$

$$f_1 = 1 + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + \left(\frac{n}{2}\right) + n + (n-1) = 4n$$

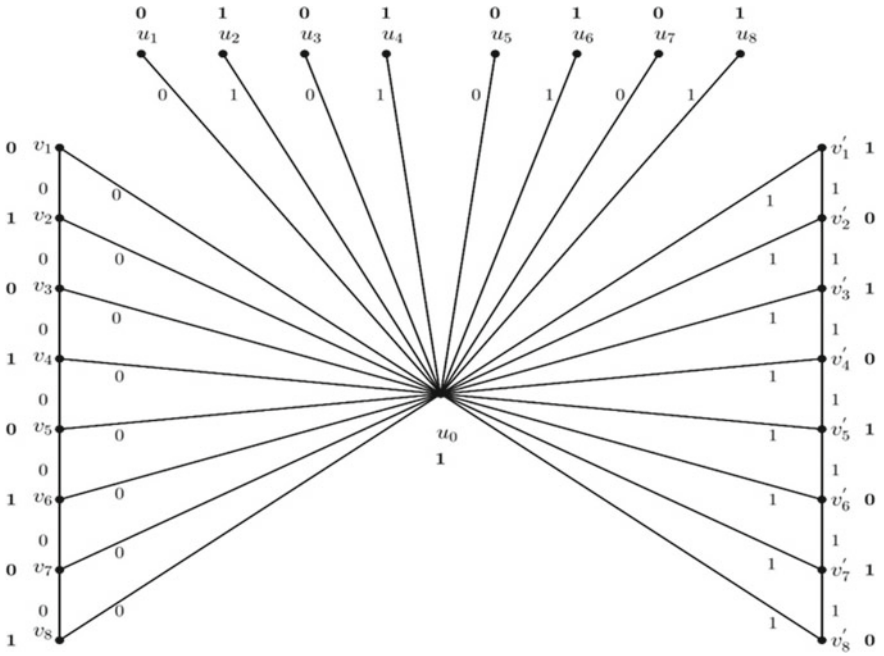


Fig. 2 Illustration of TMC labeling of V_8

$$\therefore |f_0 - f_1| = |(4n-1) - 4n| = 1$$

In both cases, it is verified that $|f_0 - f_1| \leq 1$.

Hence, vanessa graph $V_n (n \geq 2)$ acknowledges the labeling for every n.

Theorem 2.3 *Let $G = I_n$ be a graph constructed by connecting two star graph $2K_{1,n}$, $(n \geq 2)$ and two path graph $2P_n$, $(n \geq 2)$ with sharing a common node. I_n has $4n - 1$ nodes and $4n - 2$ arcs. If f: $V \cup E \to \{0, 1\}$ is a function such that $f(u) + f(v) + f(uv) = C (mod\ 2)$ for every $uv \in E(G)$ then G is total magic cordial graph.*

Proof Consider $G = I_n$ $(n \geq 2)$ the Lilly graph, where $|V| = 4n - 1$ nodes and $|E| = 4n - 2$ arcs. The node set is denoted as $V = \cup_{i=1}^{4} V_i$ where $V_1 = \{u_0\}$, $V_2 = \{u_i \text{ for } 1 \leq i \leq 2n\}$, $V_3 = \{v_i \text{ for } 1 \leq i \leq n-1\}$ and $V_4 = \left\{v'_i \text{ for } 1 \leq i \leq n-1\right\}$. The edge set is denoted as $E = \cup_{i=1}^{5} E_i$ where $E_1 = \{u_0 u_i \text{ for } 1 \leq i \leq 2n\}$, $E_2 = \{u_0 v_1\}$, $E_3 = \left\{u_0 v'_1\right\}$, $E_4 = \{v_i v_{i+1} \text{ for } 1 \leq i \leq n-2\}$ and $E_5 = \left\{v'_i v'_{i+1} \text{ for } 1 \leq i \leq n-2\right\}$.

A function $f : V \cup E \to \{0, 1\}$ as follows:

$$f(u_0) = 0 \tag{16}$$

$$f(u_i) = \begin{cases} 0 & i \text{ is odd} \\ 1 & i \text{ is even} \end{cases} \tag{17}$$

$$f(v_i) = 1 \quad \text{for } 1 \le i \le n-1 \tag{18}$$

$$f\left(v_i'\right) = 0 \quad \text{for } 1 \le i \le n-1 \tag{19}$$

$$f(u_0 u_i) = \begin{cases} 0 & i \text{ is odd} \\ 1 & i \text{ is even} \end{cases} \tag{20}$$

$$f(u_0 v_1) = 1 \tag{21}$$

$$f\left(u_0 v_1'\right) = 0 \tag{22}$$

$$f(v_i v_{i+1}) = 1 \quad \text{for } 1 \le i \le n-2 \tag{23}$$

$$f\left(v_i' v_{i+1}'\right) = 0 \quad \text{for } 1 \le i \le n-2 \tag{24}$$

To prove that $G = I_n (n \ge 2)$ posses total magic cordial labeling, we require that the difference between f_0 and f_1 is at most 1. That is, to prove that $|f_0 - f_1| \le 1$ (Fig. 3).

$$f_0 = 1 + n + (n-1) + n + 1 + (n-2) = 4n - 1$$

$$f_1 = n + (n-1) + n + 1 + (n-2) = 4n - 2$$

$$\therefore |f_0 - f_1| = |(4n-1) - (4n-2)| = 1$$

It is verified that $|f_0 - f_1| \le 1$.
Hence, lilly graph $I_n (n \ge 2)$ admits the labeling for every n.
In both the cases, it is verified that $|f_0 - f_1| \le 1$.
Hence, Lilly graph $I_n (n \ge 2)$ owns up the labeling for every n.

Theorem 2.4 *Let $G = Lo_n$ be a graph constructed from a shell graph adding a node in between each pair of adjacent nodes on the cycle and adding an arc in apex and two or more chords with $4n + 9$ nodes and $6n + 11$ arcs. If $f: V \cup E \to \{0, 1\}$ be a function such that $f(u) + f(v) + f(uv) = C \pmod 2$ for every $uv \in E(G)$ then G is total magic cordial graph.*

Proof $G = \mathrm{Lo}_n$ is lotus graph with $|V| = 4n + 9$ nodes and $|E| = 6n + 11$ arcs, where $V = V_1$, $V_1 = \{u_i \text{ for } 1 \le i \le n\}$, and $E = \cup_{i=1}^{n} E_i$ where $E_1 = \{u_1 u_2\}$,

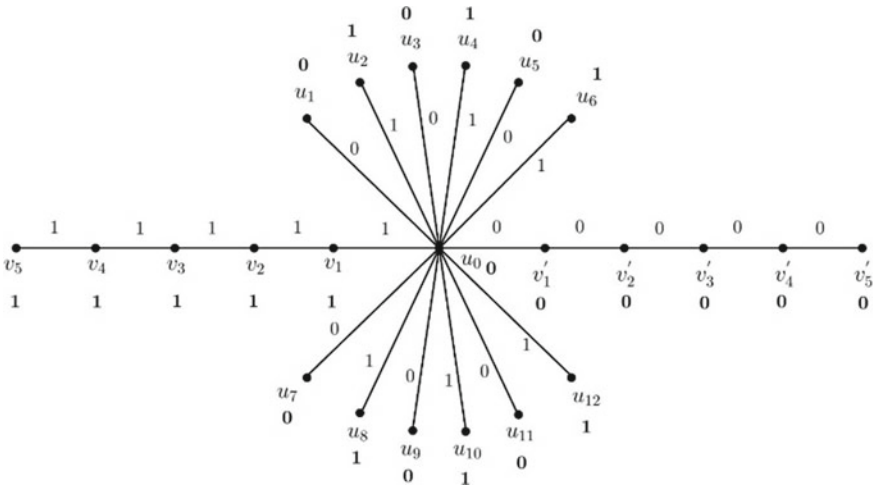


Fig. 3 Illustration of TMC labeling of I_6

$E_2 = \left\{u_2u_{2i+1} \text{ for } 1 \le i \le \left(\frac{n-1}{2}\right)\right\}$, $E_3 = \left\{u_{i+1}u_{i+2} \text{ for } 2 \le i \le \left(\frac{n-1}{2}\right)\right\}$ and $E_4 = \left\{u_iu_{i+1} \text{ for} \left(\frac{n-1}{2}\right) + 2 \le i \le n-1\right\}$. The function $f : V \cup E \to \{0, 1\}$ as follows:

$$f(u_i) = \begin{cases} 1 & i \text{ is odd} \\ 0 & i \text{ is even} \end{cases} \tag{25}$$

$$f(u_1u_2) = 0 \tag{26}$$

$$f(u_2u_{2i+1}) = \begin{cases} 0 \text{ for } 1 \le i \le \left(\frac{n-1}{4}\right) \\ 1 \text{ for } \left(\frac{n-1}{4}\right) + 1 \le i \le \left(\frac{n-1}{2}\right) \end{cases} \tag{27}$$

$$f(u_{i+1}u_{i+2}) = 1 \quad \text{for } 2 \le i \le \left(\frac{n-1}{2}\right) \tag{28}$$

$$f(u_iu_{i+1}) = 0 \quad \text{for} \left(\frac{n-1}{2}\right) + 2 \le i \le n-1 \tag{29}$$

To prove that $G = \text{Lo}_n$ confess the labeling, we need to show that, difference between f_0 and f_1 is at most by 1. That is, to prove that $|f_0 - f_1| \le 1$.

$$f_0 = \left(\frac{n-1}{2}\right) + 1 + \left(\frac{n-1}{4}\right) + \left[\left(\frac{n-1}{2}\right) - 1\right] = \left(\frac{5n-5}{4}\right)$$

$$f_1 = \left(\frac{n+1}{2}\right) + \left(\frac{n-1}{4}\right) + \left[\left(\frac{n-1}{2}\right) - 1\right] = \left(\frac{5n-5}{4}\right)$$

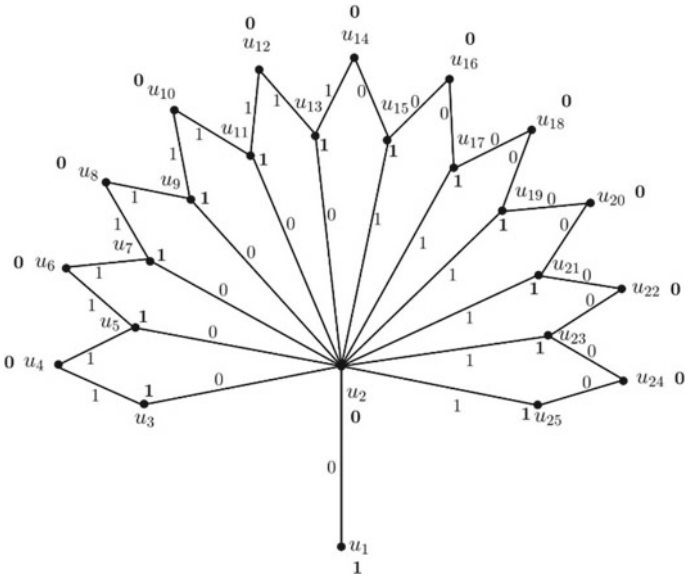


Fig. 4 Illustration of TMC labeling of Lo_{25}

$$\therefore |f_0 - f_1| = \left| \left(\frac{5n-5}{4} \right) - \left(\frac{5n-5}{4} \right) \right| = 0$$

It is verified that $|f_0 - f_1| \leq 1$.
Hence, lotus graph Lo_n owns up the labeling for every n (Fig. 4).

3 Conclusion

This article is evidence for the presence of labeling for flower graphs, namely, octopus graph, vanessa graph, graph, lilly graph, tortoise graph, and lotus graph. It was inspiring to investigate total magic cordial labeling on flower graphs. This labeling can be extended to various types of graphs and examined in the future.

References

1. Harary F (1969) Graph theory. Addison Wesley, Reading, Massachusetts
2. Gallian JA (2021) A dynamic survey of graph labeling. Electron J Comb.#DS6
3. Cahit I (1987) Cordial graphs: a weaker version of graceful and harmonious graphs. Ars Combin 23:201–208
4. Cahit I (2002) Some totally modular cordial graphs. Discuss Math Graph Theor 22:247–258

5. Parameswari R, Rajeswari R (2013) Total bimagic labeling and total magic cordial labeling of Paley digraphs. In: IEEE Conference, PRIME 2013
6. Parameswari R, Rajeswari R (2015) Total magic cordial labeling of (P_n, K_n), gear and shadow graphs. Int J Pure Appl Math 101(6):993–1001
7. Parameswari R (2015) Total magic cordial labeling of square and shadow graphs. Appl Math Sci 9(25):1229–1234
8. Jeyanthi P, Angel Benseera N, Immaculate Mary M (2013) On totally magic cordial labeling. SUT J Math 49(1):13–18

Missing Value Imputation Using Weighted KNN and Genetic Algorithm

Vikesh Kumar Gond, Aditya Dubey, Akhtar Rasool, and Nilay Khare

Abstract Missing data cause many challenges for imputing missing value in real-world datasets. Much research has been done on these challenges, but most existing research focuses on the classification task. Only a few methods can handle the large datasets for imputation of missing value. This paper proposes a new hybrid approach to impute the missing value. This approach is based on a Weighted k-nearest neighbor (WKNN) and genetic programming algorithm. This approach aims to enhance the accuracy of the imputation of missing value in symbolic regression. This paper has used different datasets with a different missing ratio of data and applied the imputation model to the datasets. This approach makes exact imputation compared to other methods like Decision Tree, Genetic programming imputation, Bayesian Regression, logistic regression, WKNN, Multilayer Perceptron, Random Forest, and Support Vector Regression.

Keywords Imputation · Genetic algorithm · WKNN · Missing value · Missing completely at random · Decision tree · Missing at random

1 Introduction

Missing data reduces the performance of the prediction models because it is common in real-world problems [1, 2]. So it is necessary to impute missing values before applying any machine learning algorithm [3, 4]. Researchers in data science and data analytics are concerned about data quality [5]. At the same time, many factors such as sampling methods, training, test, and validation affect the quality of the result of a machine learning algorithm. Handling missing values effectively is a difficult task because it needs to review all instances of datasets to detect the pattern of missingness in data [6, 7]. The missing value imputation approach relies on the precision of previously estimated missing values. Genetic programming algorithm is a good algorithm that yields a method to work for a task given by the user. This

V. K. Gond (✉) · A. Dubey · A. Rasool · N. Khare
Maulana Azad National Institute of Technology, Bhopal 462003, India
e-mail: vikeshkumargond500@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_18

method creates a population size and applies methods to it until the optimal solution achieves. Genetic algorithms use genetic operators like crossover and mutation to solve problems [8, 9].

In recent times, machine learning-based imputation approaches successfully tackle the missing value problem [10]. The general concept behind this approach is to train a model on actual data and then use it to forecast all missing values of dataset characteristics. The ability to successfully deal with the problem of missing value is an essential factor for the data mining process [11]. Missing data can be handled using many techniques such as deletion, mean, median. The efficiency of these techniques is low because it does not show any changes or correlations among characteristics of datasets [12]. In the deletion method, deleting instances with missing data is a simple method, leading to the loss of vital information. Therefore, machine learning techniques are proposed to handle the missing data. Missing data are classified into three types [13–15]. (1) Missing Completely at Random (MCAR); (2) Missing at Random (MAR); (3) Missing Not at Random (MNAR).

This paper presents a new hybrid approach for missing value imputation using weighted nearest neighbor and the concepts of genetic algorithm. The results shows that the approach performs well in different real application datasets. Section 2 describes the problem of missing data. Section 3 describes the literature survey of current imputation techniques. In Sect. 4, the description of the proposed method has been made. Section 5 covers the results and discussions. The last section represents the conclusion and future work of this paper.

2 Problem Formulation

Given a dataset $Y = \{y_{ij}\}$ represent an $(P \times Q)$ matrix where variable $\{y_{ij}\}$ denotes the value of feature j for the record i. The ith record in Y is presented by $Y = \{y_{i1}, y_{i2}, y_{i3}, \ldots, y_{iQ}\}$. To represent the pattern of missing data, consider a matrix $A = \{a_{ij}\}$ exists which has the same dimension as matrix Y, where if element $\{y_{ij}\}$ is missing then $a_{ij} = 1$ and if the element is observed then $a_{ij} = 0$. The challenge of the missing value estimation issue is mainly affected by observing the missing data pattern. If the missing ratio probabilities are more than, the model's accuracy will affect, and if the missing ratio probabilities are less than, the model will give good accuracy. It is very important to properly specify the missing data pattern when developing the imputation model. Proper specification of missing data patterns affects the utility and constraints of the model. Figure 1 represents a dataset with some missing values represented by '?'. The rows represent the record, and the columns represent the features. Since there are multiple columns that contain the missing value; hence it can be concluded the dataset has multivariable missing instances.

RECORD\ FEATURE	F_1	F_2	F_3	F_4	...	F_Q
R_1	2.3	3.2	6.1	0.002	...	-6.5
R_2	5.3	2.8	?	0.006	...	?
R_3	5.7	2.08	4.1	0.008	...	-6.8
R_4	2.5	3.3	?	0.021	...	-6.1
...	...	...	...	...	...	...
R_P	?	6.4	4.9	?	...	?

Fig. 1 Example of multivariable missing values in a dataset

3 Literature Survey

González-Vidal et al. proposed a method for missing value imputation called Bayesian Maximum Entropy (BME) [13]. BME imputes missing values in IoT environments. BME uses the spatiotemporal properties of data to estimate missing data. The data is too varied in IoT, so it is slightly complex to impute the missing data in IoT. BME estimates missing data in various IoT fields by employing low and high-accuracy sensors. BME model uses three stages Knowledge acquisition, integration, and processing. BME employs prior knowledge as an additional limitation to lead toward more accurate spatiotemporal estimation in the prior stage. The meta-prior stage comprises a specific knowledge that contains hard and soft physical datasets where the measurement error of low precision error is employed to construct interval range. In the posterior stage, both general and specific knowledge is combined to optimize the posterior probability density function.

Zhu et al. utilized a Domain-driven Increment Model (DIM) for the imputation of missing value [16]. In this model, the authors consider the impact of the missing data estimation method might be dependent on different aspects like imputation cost, imputation order, the discriminative capability of the attribute, etc. This model aims to employ available information to impute missing value orderly and achieve minimal imputation cost and high-accuracy based on the assumption. DIM utilizes the observed data in the incomplete sample to supplement seen data for the estimation phase. First, DIM identifies the need not to impute data and predictable data to minimize the imputation cost and noise. In the second phase, DIM immediately generates a scoring algorithm to assess every missing attribute that includes missing data. Finally, DIM estimates the highest-ranking missing features first and then merges them with the existing complete datasets to get the next missing features to estimate till the estimation phase is completed.

Ma et al. proposed Parallel Outlier Detection approach based on a WKNN query approach that finds the k-nearest neighbor using the Z-order curve [17]. The Outlier

detection approach based on the KNN technique efficiently detects outliers from huge datasets. Still, many approaches cannot give good results when there are high-dimensional datasets. To emphasize the significance of features in KNN, the author proposes this method. In this method, this method uses information entropy to determine the weight of every feature. It employs the Z-order curve to encode high-dimensional data into Z-value. KNN query is a common approach for identifying outliers, and it has an impact on the effectiveness and precision of the outlier outcomes. The weighted KNN of every entity is examined based on its Z-value. This means the newly outlier detection approach is signified based on the lowest and the average distance between every entity and its weighted KNN.

Jiacheng L. et al. proposed a method called a hybrid genetic approach that is basically based on information entropy and game theory [18]. This method is proposed to overcome the problem of traditional genetic algorithms. The main problem of conventional genetic algorithms is that it often falls into local optima. In a hybrid algorithm, as per the information entropy species, the diversity of the initial population is calculated utilizing the standard genetic algorithm and partial genetic algorithm. Based on the information entropy and fitness value of every subgroup population, a complete game theory operation implements to get the optimum solution for the whole population. The genetic algorithm problem can be presented in many ways. Mir et al. suggested combining the Pareto concept and genetic programming algorithm theory to solve the optimization problem by searching, optimizing, and machine learning in a genetic algorithm [19]. The author suggested the cell genetic algorithm for real number coding to resolve the optimization problem, and the imputed values were exact compared to existing methods.

4 Proposed Technique

There are various imputation techniques like MICE, in which batch imputation is done by using the available value to impute the missing value in a dataset. But this technique does not work for imputing the single test instances [20]. To overcome this problem, training data is required. This paper proposes an imputation technique based on the Weighted KNN and genetic programming algorithm (WKNN-GP). This approach gives a good result without reusing the training data for a record. The goal of the training phase is to provide an imputation model. This technique is basically divided into two parts. In the first part, incomplete training data is used for training by WKNN-GP, and it generates imputed training data. After imputed training data, an imputation model applies to it. WKNN-GP test algorithm uses incomplete test data and generates imputed test data. And then, Relative Square Error (RSE) is computed between the complete and imputed datasets (Fig. 2).

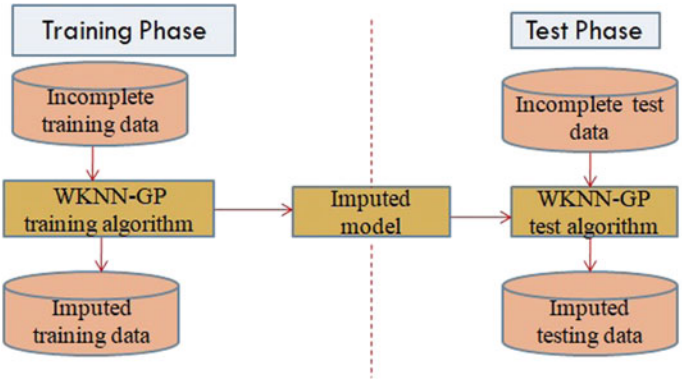


Fig. 2 Block diagram of proposed technique

5 Result and Discussion

5.1 Dataset Description

In this paper, seven different real-world datasets have been used to evaluate the model's performance. Instances represent the number of rows, and features represent the number of columns in each dataset. Initially, the dataset has none of its entries missing, so some percentage of the data is deleted for testing the proposed technique's accuracy compared to the previously available approach. The missing ratio is 10, 30, and 50%. The missing values have a multivariable missing data framework. The dataset Boston Housing, Yacht Hydrodynamics, Forest Fires, Concrete Compressive Strength, Energy Efficiency, and Weather Izmir are taken from the UCI dataset repository [21], whereas the last dataset, Kin8nm, is from the OpenML repository [22] (Table 1).

Table 1 Dataset Description

S. No.	Dataset	Instances	Features
1	Boston housing	506	14
2	Yacht hydrodynamics	308	7
3	Forest fires	517	13
4	Concrete compressive strength	1030	9
5	Energy efficiency	768	8
6	Weather izmir	1461	9
7	Kin8nm	8191	9

5.2 Performance and Evaluation

The RSE is used to measure the performance of missing data imputation [20]. Similar to Mean Square Error, and Root Mean Square Error, RSE is calculated by squaring the error residuals of the predictions. As a result, RSE is sensitive to the outlier. The design of this work is verified by objective evaluation, especially utilizing the RSE measure. The RSE measure's typical values are in the range of [0, 1]. Where the value '0' of RSE shows that the model perfectly matches the required goal, but the value '1' of RSE shows that model predicts worse than the utilizing mean prediction as a prediction model.

$$\mathrm{RSE} = \frac{\sum_{i=1}^{n} (x_i - p_i)^2}{\sum_{i=1}^{n} (p_i - \bar{p})^2} \tag{1}$$

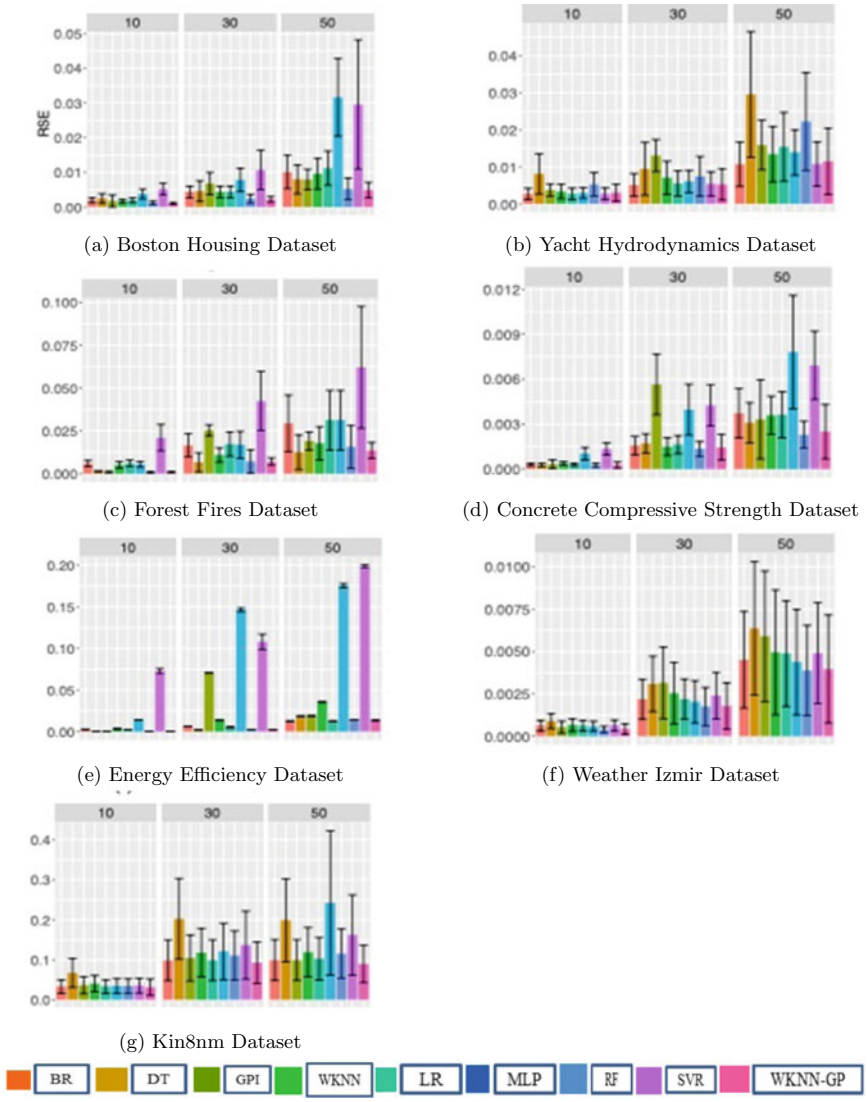
where n is the number of instances and x_i is the ith imputed values, p_i is the ith desired value, and the average of desired value p_i is denoted by $\bar{p}$, where $i = 1, 2, 3, \ldots n$.

5.3 Experimental Design

The accuracy of the model is computed for the different missing ratios. The performance of varying imputation techniques is shown in Fig. 3. The figure shows that if the missing probability is lesser, there is no massive difference in the RSE of imputation techniques. Each subfigure shows the performance comparison for different imputation techniques. The x-axis presents the missingness probability ratio of the dataset, and the y-axis presents the RSE for various imputation techniques on different datasets. By using the proposed WKNN-GP technique, 10, 30, and 50% missing ratio of Boston Housing, Yacht Hydrodynamics, Forest Fires, Concrete Compressive Strength, Energy Efficiency, Weather Izmir, and Kin8nm to have good accuracy as compared to other methods like Bayesian Regression (BR), Decision Tree (DT), Genetic programming imputation (GPI), Weighted k Nearest Neighbor (WKNN), Logistic Regression (LR), Multilayer Perceptron (MLP), Random Forest (RF), Support Vector Regression (SVR).

6 Conclusion

This paper proposes a new hybrid method for imputation of missing value based on WKNN-GP. This model is applied to different datasets like Boston Housing, Yacht Hydrodynamics, Forest Fires, Concrete Compressive Strength, Energy Efficiency, Weather Izmir, and Kin8nm. The model gave a good performance compared to other methods like Decision Tree, Genetic programming imputation, Bayesian Regression, logistic regression, WKNN, Multilayer Perceptron, Random Forest, and Support



(a) Boston Housing Dataset

(b) Yacht Hydrodynamics Dataset

(c) Forest Fires Dataset

(d) Concrete Compressive Strength Dataset

(e) Energy Efficiency Dataset

(f) Weather Izmir Dataset

(g) Kin8nm Dataset

Fig. 3 Performance comparison by RSE of different imputation technique with varying missing ratio of 10, 30, and 50% of different datasets

Vector Regression. The performance is measured on the real-world dataset using imputation accuracy. The percentage of the missing ratio of data is a crucial part of the imputation of missing values. In this paper, 10, 30, and 50% of the missing ratio of data have been taken. The experimental results clearly show that the proposed technique outperformed better accuracy and low RSE than other techniques like DT, GPI, WKNN, and SVR techniques. In the future, there is a need to do work for imputing the missing values using different missingness mechanism type data like MCAR and MNAR.

References

1. Zhao J, Nie Y, Ni S, Sun X (2020) Traffic data imputation and prediction: an efficient realization of deep learning. IEEE Access 46713–46722
2. Niemela M, Ayramo S, Karkkainen T (2021) Toolbox for distance estimation and cluster validation on data with missing values. IEEE Access, pp 352–67
3. Dubey A, Rasool A (2019) Data mining based handling missing data. In: 2019 Third international conference on I-SMAC (IoT in Social, Mobile, Analytics, and Cloud) (I-SMAC), pp 483–489
4. Dubey A, Rasool A (2000) Time series missing value prediction: algorithms and applications. In: International conference on information, communication and computing technology, pp 21–36
5. Jadhav A, Pramod D, Ramanathan K (2019) Comparison of performance of data imputation methods for numeric dataset. Appl Artif Intell 33(10):913–33
6. Ma Q, Gu Y, Lee WC, Yu G (2018) Order-sensitive imputation for clustered missing values. IEEE Trans Knowl Data Eng 31(1):166–80
7. Dubey A, Rasool A (2020) Clustering-based hybrid approach for multivariate missing data imputation. Int J Adv Comput Sci Appl (IJACSA) 11(11):710–714
8. Langazane SN, Saha AK (2022) Effects of particle swarm optimization and genetic algorithm control parameters on overcurrent relay selectivity and speed. IEEE Access 10:4550–4567
9. Singh GD, Prateek M, Kumar S, Verma M, Singh D, Lee HN (2022) Hybrid genetic firefly algorithm-based routing protocol for VANETs. IEEE Access 10:9142–9151
10. Fazakis N, Kostopoulos G, Kotsiantis S, Mporas I (2020) Iterative robust semisupervised missing data imputation. IEEE Access 8:90555–90569
11. Dubey A, Rasool A (2020) Local similarity-based approach for multivariate missing data imputation. Int J Adv Sci Technol 29(06):9208–9215
12. Duan S, Luo H, Liu H (2022) A multi-strategy seeker optimization algorithm for optimization constrained engineering problems. IEEE Access 10:7165–95
13. González-Vidal A, Rathore P, Rao AS, Mendoza-Bernal J, Palaniswami M, Skarmeta-Gómez AF (2020) Missing data imputation with Bayesian maximum entropy for internet of things applications. IEEE Internet of Things J 8(21):16108–16120
14. Yadav A, Dubey A, Rasool A, Khare N (2021) Data mining based imputation techniques to handle missing values in gene expressed dataset. Int J Eng Trends Technol 69(9):242–250
15. Gond VK, Dubey A, Rasool A (2021) A survey of machine learning-based approaches for missing value imputation. In: 2021 third international conference on inventive research in computing applications (ICIRCA), pp 841–846
16. Zhu X, Yang J, Zhang C, Zhang S (2019) Efficient utilization of missing data in cost-sensitive learning. IEEE Trans Knowl Data Eng 33(6):2425–2436
17. Ma Y, Zhao X (2021) POD: a parallel outlier detection algorithm using weighted KNN. IEEE Access 9:81765–8177

18. Jiacheng L, Lei L (2020) A hybrid genetic algorithm based on information entropy and game theory. IEEE Access 8:36602–36611
19. Mir AA, Rafique M, Hussain L, Almasoud AS, Alajmi M, Al-Wesabi FN, Hilal AM (2022) An improved imputation method for accurate prediction of imputed dataset based radon time series. IEEE Access 10:20590–20601
20. Dubey A, Rasool A (2021) Efficient technique of microarray missing data imputation using clustering and weighted nearest neighbour. Sci Rep 11(1):1–12
21. https://archive.ics.uci.edu/ml/datasets.php
22. https://www.openml.org/search?type=data

Aqua Pura: An IoT-Based System to Make Rainwater Salvageable and Manageable at a Low Cost

Naser Abdullah Alam, Anika Tahsin Momo, Sumyea Nazifa Aurpita, Nigar Sultana Anni, and Bijan Paul

Abstract In a densely populated country like Bangladesh, the groundwater levels have dramatically decreased in the past few decades while arsenic concentrations in them are increasing. In many cases, a significant amount of water is wasted due to overflow, resulting in a loss of valuable water resources as well as money. Water conservation has become more important than ever, while alternative water sources must be considered. To solve this, it is best to monitor the level of water under consideration by using IoT alongside collecting rainwater. This can help reduce the stress on groundwater in a country like Bangladesh, where we receive an annual average of 2200 mm of rainfall every year. The main obstacle, however, is that due to heavy air pollution, an undesirably high concentration of gases such as sulfur dioxide and nitrogen dioxide mixes with rainwater and causes acid rain. This research aims to achieve sustainable development by evaluating rainwater harvesting and ensuring the collection is safe for use.

Keywords Water level · Automatic pump · Rainwater collector · Rainwater harvesting · pH sensor · Water quality · Water conservation · Water sustainability · Groundwater · Acid rain · Water degradation

N. A. Alam · A. T. Momo · S. N. Aurpita · N. S. Anni · B. Paul (✉)
University of Liberal Arts Bangladesh, Dhaka, Bangladesh
e-mail: bijan.paul@ulab.edu.bd

N. A. Alam
e-mail: naser.abdullah.cse@ulab.edu.bd

A. T. Momo
e-mail: anika.tahsin2.cse@ulab.edu.bd

S. N. Aurpita
e-mail: sumyea.nazifa.cse@ulab.edu.bd

N. S. Anni
e-mail: nigar.sultana.cse@ulab.edu.bd

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_19

1 Introduction

The dramatic decrease in groundwater levels has been one of the major concerns for the last three decades. Due to unmonitored extraction and careless use, the water level has dropped as low as sixty meters in some areas [1, 2]. To conserve the underground water supply, we have and look for alternative sources of water. Bangladesh receives an annual average of 2200 mm of rainfall every year. However, due to excessive air pollution, the density of toxic gases such as sulfur dioxide and nitrogen oxide has increased drastically [3] (Fig. 1).

The high concentrations of sulfur dioxide and nitrogen oxide react with rainwater, producing sulfuric and nitric acid, which results in acid rain. Bangladesh has a high population density, which together makes harvesting rainwater and directly using it for general purposes impossible [3–5] due to the land area shortage and high demand [6].

This system was developed to conserve water and use rainwater as an alternate source of water supply, also separating them by their pH levels. The system collects rainwater and monitors its pH. If the sensor identifies the water as acidic, it is passed on to the open drain waste system. If the pH of the collected water is neutral, then it is pumped to the mesh filter to filter out any debris and channeled into the main water tank. This system can also control and indicate the water level in a tank. This system enables it to automatically turn on a water pump when the water level gets low, refill the water back to an adequate level, and turn off the pump when the tank is full. Afterward, the pH level and water level will be shown on the LCD monitor.

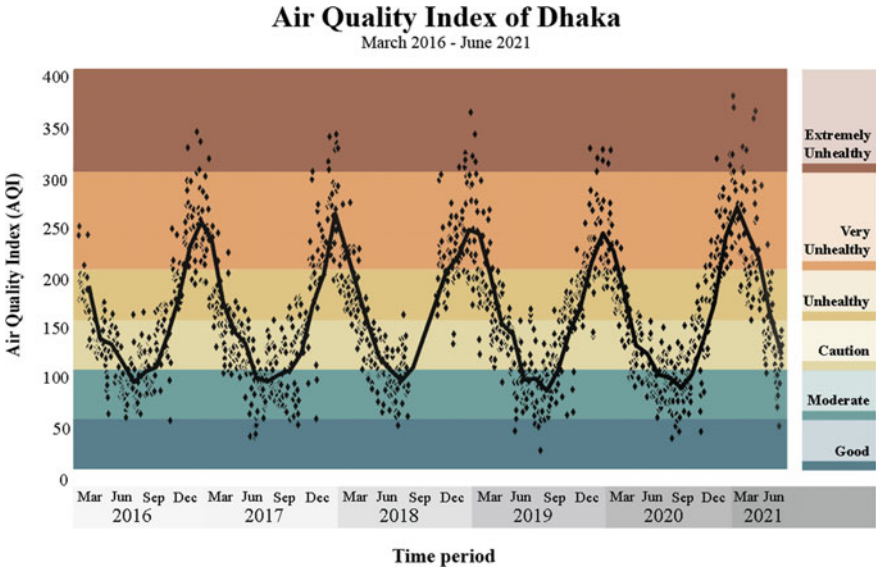


Fig. 1 Dhaka Air Quality Index (AQI) from the year 2016 to 2021. *Source* U.S Embassy Bangladesh

2 Related Work

There are a few existing systems for rainwater harvesting. However, the collected rainwater has no direct interaction with users [7]. The ones that consider direct interaction with users require a completely new water line [8], or use components that are not waterproof, such as servo motors to control the water flow after monitoring its quality [9], or collect wastewater from regular household use, such as wastewater from the kitchen, or the shower, or washing machine. These systems focus more on checking water turbidity than pH [10, 11].

3 Applications

1. Anywhere where there is rain and a need to use water.
2. Urban Residential Areas—Safe-checking rooftop rainwater collection to ease pressure on WASA as well as reduce flooding in city.
3. Urban Residential Water Management—Can synchronize rain reserve and WASA line to make efficient the use of water.
4. Rural Areas—Collection of safe rainwater to avoid health hazards, especially to children and unaware populace.
5. Farms—Can ensure safe environment in algae farms, fisheries food production, etc.
6. Industrial Areas—Measuring factories contribution to acid rain, increasing awareness and corporate responsibility, will encourage factories to shift to greener management of waste.

4 Working Mechanism

Our system works as both a water level management system and a rainwater harvesting system. Underground water is collected and stored in the main water reserve until there is no more rainwater stored in the rainwater reserve. Once rainwater has been collected and separated by its pH levels, it will either be stored in the main water reserve or be passed to an open drainage system. Before the water is stored or discharged, it is filtered by a mesh filter to remove any slag and debris and discharge them to the open drain waste system. After that, as the water level decreases in the main water reserve, the water is pumped into it from either the underground supply or the rainwater reserve (Fig. 2).

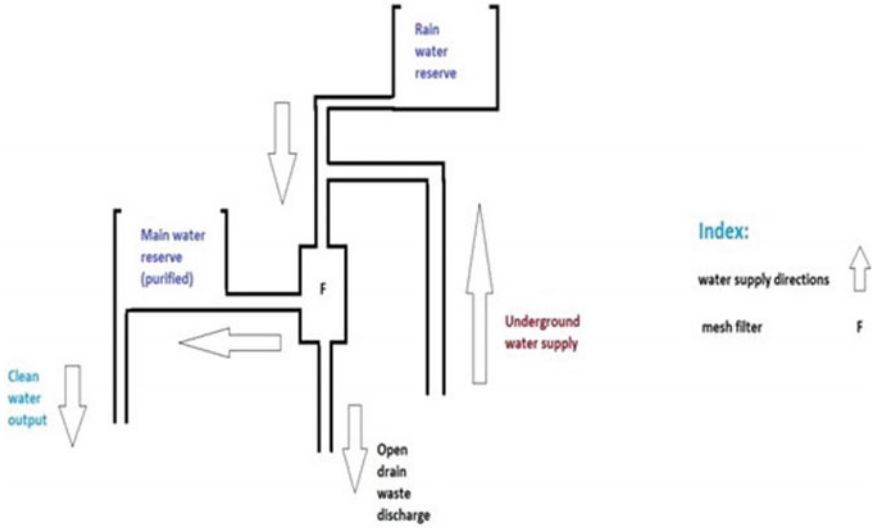


Fig. 2 Working mechanism of the system

5 Programming Mechanism

Using the sonar sensor, we first detect the water level in the rainwater storage. If the water level is the same as the water tank height, then the pH level is not checked, but if it is less than the water tank height, the sensor first checks the pH level of the stored water. If the pH of the stored water is less than three, it is discharged through the drainage using pump one, but if it is more than three, we pass the stored water into the main water storage using pump two (Fig. 3).

6 Circuit Diagram

The connections used in the system are depicted in the circuit diagram above. As can be seen, the system is built on an Arduino Uno board, but it may also be built on an Arduino Nano board. The pH sensor's output pin is attached to the Arduino Uno's analog pin A0. The i2C module for the 16 × 2 LCD is attached to the analog pins A4 and A5. The digital pin D2 is connected to relay module 1 (RM1), which controls water pump 1 (WP1), which redirects the water flow when the pH is below three (acidic). The digital pin D3 is connected to relay module 2 (RM2), which controls water pump 2 (WP2), which redirects the water flow when the pH is above three (normal rainwater). The digital pins D6 and D7 are connected to the sonar sensor; one sends the triggering frequency and the other receives and sends the data to the

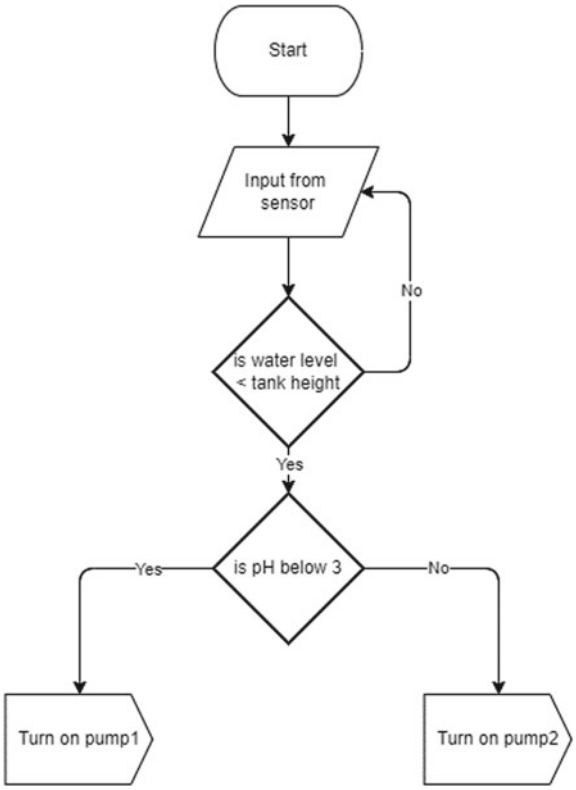


Fig. 3 Programming mechanism of the system

Arduino. All the sensors and components are connected to a common ground and a common VCC, the 5 V 3A DC adapter (Fig. 4).

7 Required Hardware Components

See Table 1.

8 Design and Implementation

As shown in the figure above, the device can successfully detect the pH and water level in the water tank. Based on the pH and water level, the device controls two 5 V relays that can be used to control the water pumps and redirect the water based on pH levels. In the system, the red bicker holds the acid water (rain). The blue bicker holds useable water. These relays are capable of channeling 220 V electricity, so the

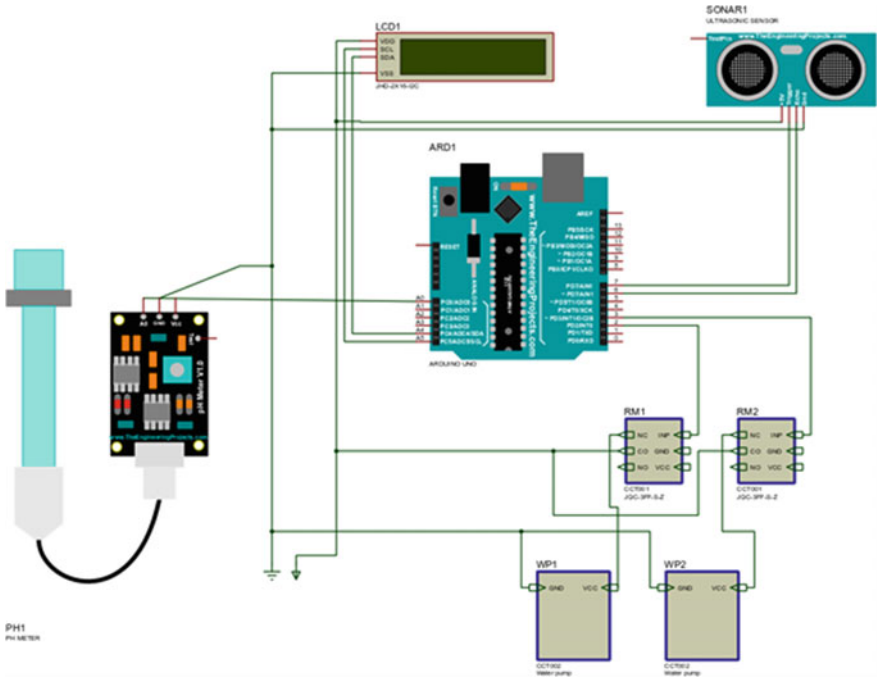


Fig. 4 Circuit diagram of the system

Table 1 Hardware components required for the system

Products
1.Arduino Uno or Arduino Nano
2. pH sensor with BNC connector module
3. Sonar sensor
4. 5 V Single channel relay module
5. 5 V DC water pump
6. LCD display with I2C module
7. 5 V 3A DC wall power adapter

relays can be directly connected to 220 V AC pumps. The system's production was completed over the course of seven days (Fig. 5).

9 Future Prospects and Conclusion

In the future, we plan to upgrade this system by adding Internet support [12] and a limestone filter that will neutralize the acidic water and make it reusable [13]. Our

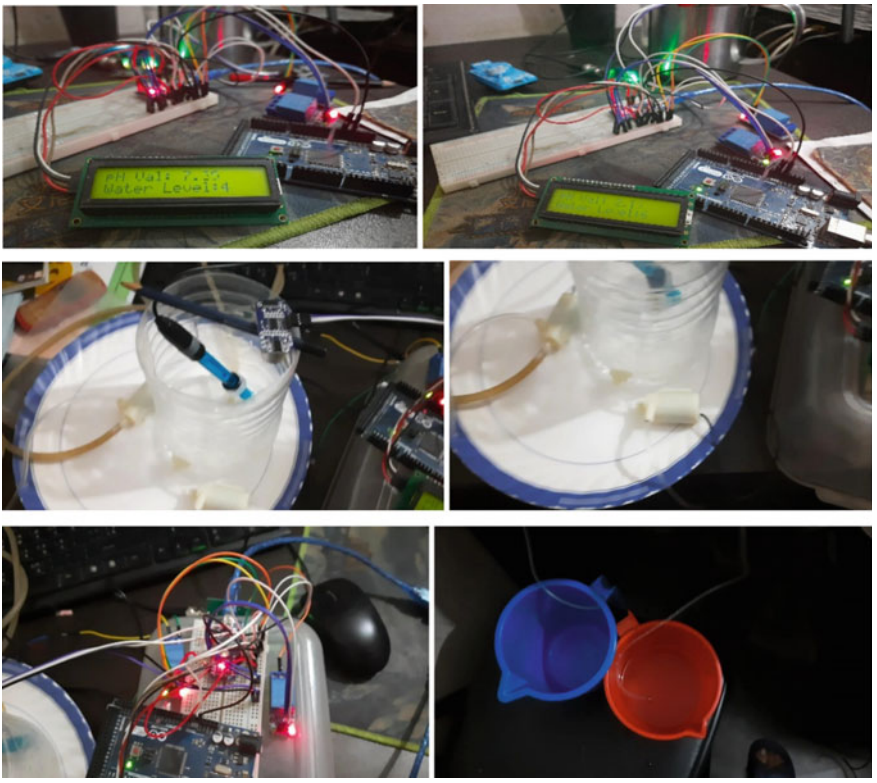


Fig. 5 Design and implementation of the system

system can also be configured to detect the level of water in the main water tank and control the main AC 220 V water pump [14].

Our system employs a simple mechanism to detect and indicate the water level and purity in any water container. The system's operations will be necessary in rural, urban, and industrial areas. It efficiently utilizes multiple available water sources. It is a simple yet effective way to prevent the wastage of water. Its simplicity in design and low-cost components makes it a very innovative solution to solve water resource problems. Its versatility means it could be used in industrial sectors and cut off acid rain in the first place.

References

1. Mamoon W, Quadir A, Mohib K, Munir M, Haque A, Billah M (2020) Analysis of groundwater level fluctuations in Dhaka city. In: 5th International conference on advances in civil engineering (ICACE-2020), CUET, Chattogram, Bangladesh, Dec 2020

2. Hoque MA, Hoque MM, Ahmed KM (2007) Declining groundwater level and aquifer dewatering in Dhaka metropolitan area, Bangladesh: causes and quantification. Hydrogeol J 15(8):1523–1534. https://doi.org/10.1007/s10040-007-0226-5
3. Rahman MM, Mahamud S, Thurston GD (2019) Recent spatial gradients and time trends in Dhaka, Bangladesh, air pollution and their human health implications. J Air Waste Manag Assoc 69(4):478–501. https://doi.org/10.1080/10962247.2018.1548388
4. Rahman S, Khan MTR, Akib S, Din NBC, Biswas SK, Shirazi SM (2014) Sustainability of rainwater harvesting system in terms of water quality. Sci World J 2014:1–10. https://doi.org/10.1155/2014/721357
5. Sarkar A, Rahman AK, Samad A, Bhowmick A, Islam J (2019) Surface and ground water pollution in Bangladesh: a review. 6:47–69. https://doi.org/10.20448/journal.506.2019.61.47.69
6. Haque MM, Rahman A, Samali B (2016) Evaluation of climate change impacts on rainwater harvesting. J Clean Prod 137:60–69. https://doi.org/10.1016/j.jclepro.2016.07.038
7. Ward S, Memon FA, Butler D (2012) Performance of a large building rainwater harvesting system. Water Res 46(16):5127–5134. https://doi.org/10.1016/j.watres.2012.06.043
8. Mun JS, Han MY (2012) Design and operational parameters of a rooftop rainwater harvesting system: definition, sensitivity and verification. J Environ Manage 93(1):147–153. https://doi.org/10.1016/j.jenvman.2011.08.024
9. Ranjan V, Reddy MV, Irshad M, Joshi N (2020) The internet of things (IOT) based smart rain water harvesting system. IEEE Xplore. https://doi.org/10.1109/ICSC48311.2020.9182767
10. Verma G, Gupta S, Gupta R (2020) IoT based technique for household rainwater harvesting. IEEE Xplore. https://doi.org/10.1109/GCWOT49901.2020.9391634
11. Verma G (2021) Low cost smart ground system for rainwater harvesting for Indian houses using IoT technology. Res Square. https://doi.org/10.21203/rs.3.rs-531034/v1
12. iCircuit (2016) Interfacing arduino uno with nodemcu. Retrieved from: https://icircuit.net/interfacing-arduino-uno-wtih-nodemcu/1474#:~:text=Connecting%20nodemcu%20with%20arduino%20uno%20Once%20you%20verified,SoftSerial%20Tx%20to%20NodeMCU%20Rx%20%28through%20level%20shifter%29
13. Abusel L, Karabsheh R, Edor Y, Shariah R (2020) Neutralizing acid rain in fresh waters by using alkaline. J Univ Stud Inclusive Res 1(2)
14. Godwin Premi MS, Malakar J (2019) Automatic water tank level and pump control system. In: 2019 International conference on intelligent computing and control systems (2019). https://doi.org/10.1109/iccs45141.2019.9065438

FMDCDTL: Design of Fusion-Based Model for Identification of Drowsy Drivers via Cascaded Deep Transfer Learning

Rashmi A. Wakode and S. W. Mohod

Abstract Identification of drowsiness while driving requires continuous driver monitoring, and wake-up alert modelling, which assists in reducing on-road accidents. To perform this task, multimodal analysis including, eye tracking, yawn tracking, oxygen level monitoring, road parameters, etc. are evaluated, and machine learning models are applied to them for classification of drowsiness conditions. But most of these models have limited accuracy and precision performance when applied to real-time conditions. This limitation arises, because the models are trained on clinical datasets, which lack in feature density under real-time environments. To overcome this limitation, a novel fusion-based model for identification of drowsy drivers via cascaded deep transfer learning is proposed in this text. The proposed model initially uses standard datasets for eye tracking, yawn tracking, electroencephalogram (EEG), and driving behaviour to train an ensemble learning model, which is fine tuned via on-road conditions. Data collected from on-road conditions is processed via a correlation engine which assists in similarity identification and transfer learning processes. The proposed model uses mel frequency cepstral coefficients (MFCC) with i-Vector features and classifies them via augmentation of k-nearest neighbours (kNN), random forest (RF), linear support vector machine (LSVM), logistic regression (LR), multilayer perceptron (MLP), and customized 1D convolutional neural network (CNN) classifiers. Results of the model were evaluated in terms of accuracy, precision, recall, f-Measure, and response time metrics, and were compared with various state-of-the-art methods. It was observed that the proposed model showcased 8.3% better accuracy, 6.5% better precision, 3.4% better recall, and 14.5% faster response than existing models. Due to this performance enhancement, the proposed model is deployable for a wide variety of real-time application scenarios.

Keywords Drowsy · Multimodal · Features · MFCC · i-Vector · kNN · SVM · RF · MLP · 1D CNN

R. A. Wakode (✉) · S. W. Mohod
Prof. Ram Meghe Institute of Technology and Research, Badnera, Amravati, India
e-mail: rawakode@mitra.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_20

1 Introduction

Design of a driver drowsiness classification model involves multi-domain inputs that include driver's facial analysis, mood analysis, physiological parameter analysis, driving pattern analysis, etc. The model initially identifies face position via cascaded wavelet features, which assists in localization of facial components including eyes, nose, and mouth, along with their width and height features, which are used to optimize the region of interest (ROI) for extraction of eye pairs and mouth position values. These features along with training set classes are given to a machine learning classifier like convolutional neural network (CNN), recurrent NN (RNN), etc., for identification of drowsiness conditions. The trained model is directly used for real-time video sequences, which assists in high performance drowsiness detection under different illumination conditions.

Similar models are proposed by researchers, which assists in low error and high-efficiency drowsiness detection from real-time image sets. A review of these models [1–3] along with their nuances, advantages, limitations, and future research scopes is discussed in the next section of this text. The limitation arises like limited accuracy and precision performance, when applied to real-time conditions, because the models are trained on clinical datasets, which lack in feature density under real-time environments. To overcome these limitations, Sect. 3 discusses design of a fusion-based model for identification of drowsy drivers via cascaded deep transfer learning, wherein mel frequency cepstral coefficients (MFCC) with i-Vector features are extracted from facial, EEG, and driving pattern values. These values are classified via an augmentation of k-nearest neighbours (kNN), random forest (RF), linear support vector machine (LSVM), logistic regression (LR), multilayer perceptron (MLP), and customized 1D convolutional neural network (CNN) classifiers. Performance of the proposed model is evaluated in Sect. 4.,wherein different multimodal inputs are used for evaluation of precision, recall, accuracy, and response time for drowsiness detection under real-time conditions.

2 Literature Review

A wide variety of models are proposed for drowsy driver analysis, and each of them utilize a particular behaviour set that includes EEGs, facial patterns, driving patterns, etc. For instance, work in [1–3] proposes use of thermal imaging cameras with neural networks (TIC NN), 3D NNs, and EEG signals with percentage of eyelid closure (PERCLOS) estimation are discussed for different scenarios. The models aim at utilizing different feature vectors that can assist in low complexity, and high accuracy feature extraction and classification tasks. Similarly, facial features like eye aspect ratio (EAR) [4] and fuzzy logic with eye blinks (FL) [5] are proposed for moderate accuracy and real-time classification scenarios. But all these models have their own challenges, some of which include exclusion of real-time input patterns,

lack of physiological parameter consideration, etc. Survey of these issues is discussed in [6], which assists in covering them while design of high-efficiency drowsiness detection models. Based on this discussion, deep belief network (DBN) with facial dynamic fusion [7], single EEG with gradient boosted decision tree [8], detection of Doppler shift caused by nodding, operating steering wheel, and yawning via long short-term memory (LSTM) [9], adaptive batch-image modelling (ABI) [10], and personal-based hierarchical modelling [11] are proposed by researchers. Each of these models has their own characteristics, but are not comprehensive in nature, which limits their performance under real-time inputs.

A radio frequency identification (RFID) model that uses LSTM to identify nodding movements is proposed in [12], which assists in low delay and high accuracy drowsiness evaluation under real-time scenarios. But the model is impractical and does not cater to wider driver markets. To improve this scalability, work in [13–15] proposes use of light U-Net (LUN) for eye tracking, modelling of gamification for keeping the driver awake under real-time driving conditions, and multiple view EEG-based drowsiness detection via transfer TSK fuzzy model are discussed and validated under different applications. Similarly, work in [16–18] proposes use of EEG for eye blinks, EEG with fuzzy independent phase-locking value representations that use Lagrangian-based deep NNs (FIPL VR LDNN) and deep convolutional model based on feature recalibration are discussed by researchers for real-time application scenarios. But these models require specific hardware and has lower efficiency under actual driving conditions. To overcome this limitation, next section proposes use of fusion-based model for identification of drowsy drivers via cascaded deep transfer learning, which can be deployed for a wide variety of real-time application scenarios.

3 Design of the Proposed Fusion-Based Model

From the literature review, it was observed that existing models for drowsiness detection either utilize singular modal feature vectors, due to which their scalability and accuracy performance is limited when deployed for real-time conditions. To overcome this limitation, an improved novel fusion-based model is discussed in this text. The proposed model is depicted in Fig. 1.

From this flow, it can be observed that facial features are 2D image data and are processed via the VGGNet-19 CNN model, while driving patterns and EEG features produce 1D features, which are classified via use of ensemble classifier that combines MLP, RF, LSVM, kNN, and LR models. Results of both classification models are combined via use of a statistical analyser which assists in high-efficiency and low error classifications.

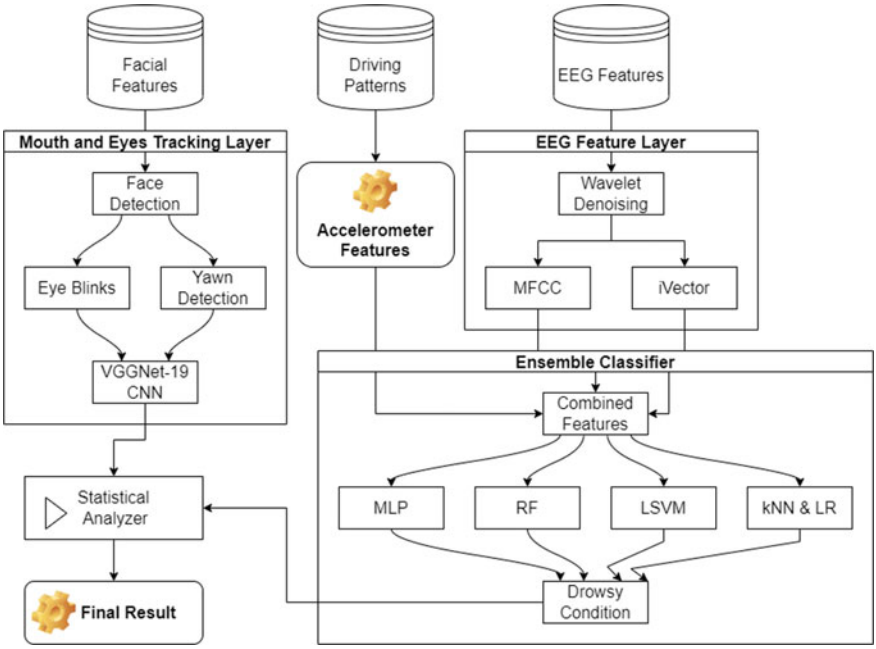


Fig. 1 Overall flow of the proposed model

3.1 Design of the VGGNet-19-Based CNN Model for Analysis of Facial Features

Real-time video sequences were captured from mobile phones and processed via local binary pattern (LBP)-based frontal face classifiers. These classifiers consist of face detection, eye localization, and mouth localization features, which assist in identification of the given facial identities. Figure 2 depicts the process of face identification from the Haar cascade features, wherein eye components, nose components, and mouth positions are identified for further classification process.

It combines edge features, line features, and four rectangle features in order to create separate identification vectors for face, eyes, and mouth locations. RoI for eyes and mouth are extracted and used to train a VGGNet-19-based CNN model, which assists in identification of eye blinks and yawning conditions.

The RoI regions of mouth and eyes are scaled to 128 $\times$ 128 and given to the VGGNet-19 model for classification into drowsy and non-drowsy categories. The convolutional layers output a large feature array.

The feature vectors have high redundancy, which can limit the classification performance for larger inputs. To avoid this condition, all convolutional features are processed via a variance maximization layer using Max Pooling operations.

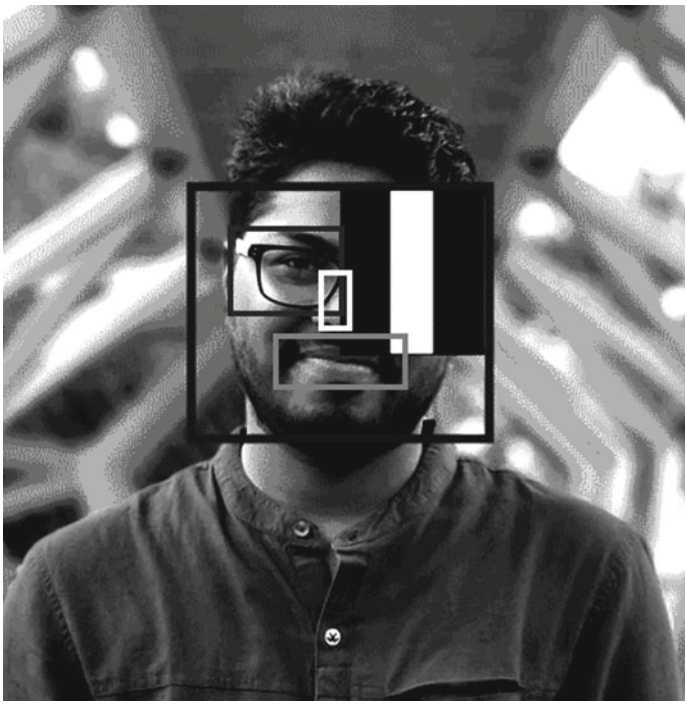


Fig. 2 Process of Haar cascade features for identification of face, eyes, nose, and mouth locations

The FCNN model uses a SoftMax-based activation function which assists in performing back-propagation-based training for continuous performance improvement. The output of the SoftMax model that indicates drowsiness (d_{out}) is evaluated via Eq. 1 as follows

$$d_{\text{out}} = \text{SoftMax}\left(\sum_{i=1}^{N_f} \text{Conv}_i * w_i + b\right) \tag{1}$$

where Conv_i represents convolutional feature vector value, w_i represents its weight value which is continuously tuned by the CNN model, and b represents a bias value that is used for correcting the output of CNN model and assists in optimization of performance. Results of this model are evaluated for eye blinks and yawning images and are used for identification of real-time drowsiness conditions. These results are combined with EEG and driving patterns to further tune performance of the underlying model for real-time scenarios. The process of driving pattern analysis and EEG classification is depicted in the next section of this text.

3.2 Design of Feature Extraction Model for Driving Patterns and EEG Signals

To identify driving patters, an accelerometer is connected on the vehicle, and its $X, Y\&Z$ values are constantly monitored and processed. Each of these values is processed via a temporal difference operator, which assists in continuous evaluation of changing driving patterns. The temporal difference operator is applied via Eq. 2, wherein $X, Y\&Z$ values are processed separately for final analysis.

$$\mathrm{TD}_{Q_i} = \mathrm{TD}_{Q_{i-1}} + \emptyset * \left[Q_{i+1} - Q_i\right] \tag{2}$$

where TD represents temporal difference, Q represents the quantity which is being processed ($X, Y or Z$), $\emptyset$ represents a scaling constant, and i represents the time instant of these evaluations.

To further process these patterns and the EEG signals for feature extraction, they are initially converted into a quantized range of (0, 1), which assists in feature normalization and assists in better classification performance. The quantization process is performed via Eq. 3 as follows

$$Q_a = \frac{N_a - \min(N_a)}{\max(N_a) - \min(N_a)} \tag{3}$$

where $Q_a, \&N_a$ represents quantized features and non-quantized features, respectively. The quantized features are converted from temporal domain to mel frequency domain. Using these components, cepstrum values are evaluated. These cepstral components are passed through a triangular activation function.

This waveform is passed through a log likelihood model, which assists in estimation of MFCC feature vectors. Similar to MFCC, i-Vector components are also evaluated, which assists in representing both EEG and driving patterns via high-density feature vectors. These i-Vectors are evaluated via Eq. 4, wherein variance calculation of features is used for extraction of final features.

$$i\mathrm{Vector}_i = \mathrm{MAX}\left(\bigcup_{j=1}^{N} x_j\right) + \begin{bmatrix} (1,1)_{\mathrm{var}} & \cdots & (1,n)_{\mathrm{var}} \\ \vdots & \ddots & \vdots \\ (n,1)_{\mathrm{var}} & \cdots & (n,n)_{\mathrm{var}} \end{bmatrix} * x_i \tag{4}$$

where N represents total number of EEG and temporal difference samples, and x represents value of that input, while $(x, y)_{\mathrm{var}}$ represents total variance levels between features x and y which is evaluated via Eq. 5 as follows

$$(x, y)_{\mathrm{var}} = \frac{\exp\left(\frac{x^2}{2}\right)}{2 * \pi * \mathrm{var}(x) * \mathrm{var}(y)} \tag{5}$$

where var(x) represents value of variance for the given input x, assists in maximizing feature variance via redundancy reduction process. All these features are given to an ensemble classification layer which is described in the next section of this text.

3.3 Design of the Ensemble Learning and Statistical Analysis Layer for Drowsiness Classification Decisions

The combined features of MFCC, i-Vector, for both EEG and driving patterns are processed via MLP, RF, LSVM, kNN, and LR models. These models are used in their original form, and final drowsy class is evaluated via Eq. 6 as follows

$$\begin{aligned} D(\text{Out})_i = \; & D(\text{MLP})_i * \text{MLP}(W) + D(\text{RF})_i * \text{RF}(W) + D(\text{LSVM})_i * \text{LSVM}(W) \\ & + D(\text{LR})_i * \text{LR}(W) + D(\text{kNN})_i * \text{kNN}(W) \end{aligned} \tag{6}$$

where $D(C)$ represents drowsiness decision by the classifier C and W represents classifier weight.

Similar to this process, the results from blink detection and yawn detection are also combined via Eq. 7 to obtain the final drowsiness state of the user

$$\begin{aligned} D(\text{Final}) = \; & D_{\text{out}}(\text{EEG}) * W(\text{EEG}) + D_{\text{out}}(\text{DP}) * W(\text{DP}) \\ & + D_{\text{out}}(\text{Eye}) * W(\text{Eye}) + D_{\text{out}}(\text{Yawn}) * W(\text{Yawn}) \end{aligned} \tag{7}$$

where $D_{\text{out}}(F)$ represents output drowsiness state via individual feature F and $W(F)$ represents weight of the feature.

4 Result Evaluation and Comparison

The proposed model uses a combination of EEG, driving patterns, Yawn detection, and blink detection in order to estimate real-time drowsy condition of drivers. To evaluate this model, accuracy, precision, recall, f-Measure, and response time values were estimated under different testing scenarios. A total of 10,000 real-time different samples of EEG, facial videos, and driving patterns were used for this evaluation and were divided into a ratio of 60:20:20, wherein 60% samples were used for training, 20% were used for testing, and remaining 20% were used for validation of the model with different test inputs. Test i-measures for different models are shown in Fig. 3.

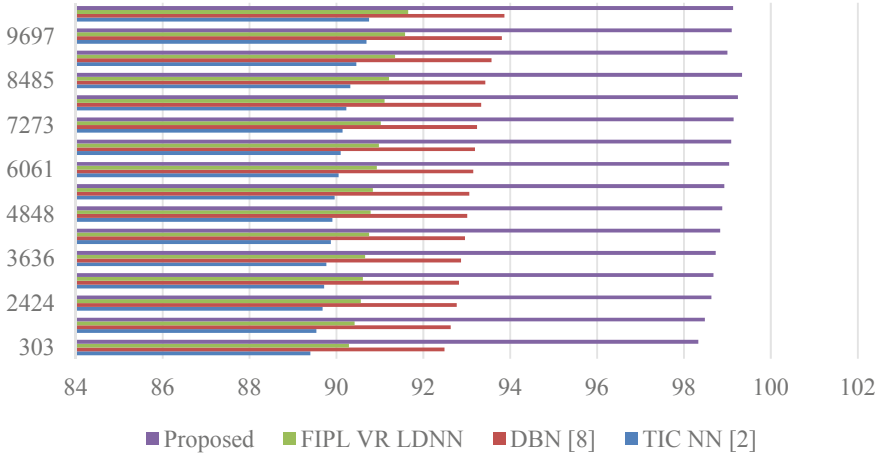


Fig. 3 Test i-Measure for different models

5 Conclusion and Future Scope

The proposed model uses a combination of EEG signals, driving patterns, and facial features in order to estimate drowsy state of drivers in real-time scenarios. Due to EEG, any changes in neurological patterns are detected, while using facial patterns (eyes and yawn), any changes in face-based features are easily tracked, and acted upon with alerts. Driving patterns are used as a post drowsiness feature vector, and alert the driver in case they are almost drowsy, due to which changes in normal driving patterns are detected under real-time conditions. These signals were represented in terms of MFCC, i-Vector, and temporal difference signals, which were classified via 2D CNN and ensemble classifiers. Due to which, the proposed model was 5.2% better than TIC NN [1], 4.6% better than DBN [7], and 4.9% better than FIPL VR LDNN in terms of accuracy under different input samples, while, it was also observed that the proposed model is 9.5% better than TIC NN [1], 5.9% better than DBN [7], and 8.3% better than FIPL VR LDNN in terms of precision under different input samples. Similar observations were done for recall and f-Measure values, which makes the model highly useful for real-time application deployments. Due to use of pre-trained classification models, the proposed method was 26.5% better than TIC NN [1], 18.4% better than DBN [7], and 23.6% better than FIPL VR LDNN in terms of delay, which assists in taking high-speed decisions and makes the model highly applicable for real-time scenarios. In future, researchers can add more parameters for drowsiness detection and can validate performance with other models. Furthermore, CNN can be replaced with recurrent NN and Q-learning-based methods, which will further assist in improving in classification performance under real-time conditions.

Table 1 Test delay for combined dataset

Number of samples	TD (ms) TICNN [1]	TD (ms) DBN [7]	TD (ms) FIPLVRLDNN	TD (ms) proposed	Number of samples	TD (ms) TICNN [1]	TD (ms) DBN [7]	TD (ms) FIPLVRLDNN	TD (ms) proposed
303	16.8	10.5	15	3.4	6061	15.5	9.1	13.6	2
1515	16.5	10.2	14.7	3.1	6667	15.3	9	13.5	1.8
2424	16.2	9.9	14.4	2.8	7273	15.3	8.9	13.4	1.8
3030	16.1	9.7	14.3	2.6	7879	15.1	8.7	13.3	1.6
3636	16	9.7	14.2	2.6	8485	14.9	8.5	13.1	1.4
4242	15.8	9.5	14	2.4	9091	14.6	8.3	12.8	1.6
4848	15.7	9.3	13.9	2.2	9697	14.2	7.8	12.4	1.3
5455	15.6	9.3	13.8	2.2	10,000	13.9	7.5	12	1.3

References

1. Kajiwara S (2021) Driver-condition detection using a thermal imaging camera and neural networks. Int J Automot Technol 22:1505–1515. https://doi.org/10.1007/s12239-021-0130-3
2. Wijnands JS, Thompson J, Nice KA et al (2020) Real-time monitoring of driver drowsiness on mobile platforms using 3D neural networks. Neural Comput Applic 32:9731–9743. https://doi.org/10.1007/s00521-019-04506-0
3. Arefnezhad S, Hamet J, Eichberger A et al (2022) Driver drowsiness estimation using EEG signals with a dynamical encoder–decoder modeling framework. Sci Rep 12:2650. https://doi.org/10.1038/s41598-022-05810-x
4. Biswal AK, Singh D, Pattanayak BK, Samanta D, Yang MH (2021) IoT-based smart alert system for drowsy driver detection. Wireless Commun Mobile Comput 2021(6627217):13. https://doi.org/10.1155/2021/6627217
5. AlKishri W, Abualkishik A, Al-Bahri M (2022) Enhanced image processing and fuzzy logic approach for optimizing driver drowsiness detection. Appl Comput Intell Soft Comput 2022(9551203):14. https://doi.org/10.1155/2022/9551203
6. Triyanti V, Iridiastadi H (2017) IOP Conf Ser: Mater Sci Eng 277:012042
7. Zhao L, Wang Z, Wang X, Liu Q (2018) Driver drowsiness detection using facial dynamic fusion information and a DBN. IET Intell Transp Syst 12:127–133. https://doi.org/10.1049/iet-its.2017.0183
8. Wang P, Min J, Hu J (2018) Ensemble classifier for driver's fatigue detection based on a single EEG channel. IET Intell Transp Syst 12:1322–1328. https://doi.org/10.1049/iet-its.2018.5290
9. Xie Y, Li F, Wu Y, Yang S, Wang Y (2021) Real-time detection for drowsy driving via acoustic sensing on smartphones. IEEE Trans Mobile Comput 20(8):2671–2685. https://doi.org/10.1109/TMC.2020.2984278
10. Kim W, Lee Y-K, Jung W-S, Yoo D, Kim D-H, Jo K-H (2020) An adaptive batch-image based driver status monitoring system on a lightweight GPU-equipped SBC. IEEE Access 8:206074–206087. https://doi.org/10.1109/ACCESS.2020.3035393
11. Chiou C, Wang W, Lu S, Huang C, Chung P, Lai Y (2020) Driver monitoring using sparse representation with part-based temporal face descriptors. IEEE Trans Intell Transp Syst 21(1):346–361. https://doi.org/10.1109/TITS.2019.2892155
12. Yang C, Wang X, Mao S (2020) Unsupervised drowsy driving detection with RFID. IEEE Trans Veh Technol 69(8):8151–8163. https://doi.org/10.1109/TVT.2020.2995835
13. Zhuang Q, Kehua Z, Wang J, Chen Q (2020) Driver Fatigue detection method based on eye states with pupil and iris segmentation. IEEE Access 8:173440–173449. https://doi.org/10.1109/ACCESS.2020.3025818
14. Ibe T, Fujiwara K, Hiraoka T, Abe E, Yamakawa T (2021) Development of game-like system using active behavior input for wakefulness-keeping support in driving. IEEE Trans Intell Vehicles 6(2):323–332. https://doi.org/10.1109/TIV.2020.3029260
15. Jiang Y, Zhang Y, Lin C, Wu D, Lin C-T (2021) EEG-based driver drowsiness estimation using an online multi-view and transfer TSK fuzzy system. IEEE Trans Intell Transp Syst 22(3):1752–1764. https://doi.org/10.1109/TITS.2020.2973673
16. Shahbakhti M et al (2022) Simultaneous eye blink characterization and elimination from low-channel prefrontal EEG signals enhances driver drowsiness detection. IEEE J Biomed Health Inform 26(3):1001–1012. https://doi.org/10.1109/JBHI.2021.3096984
17. Reddy TK, Arora V, Gupta V, Biswas R, Behera L (2022) EEG-based drowsiness detection with fuzzy independent phase-locking value representations using Lagrangian-based deep neural networks. IEEE Trans Syst, Man, Cybern: Syst 52(1):101–111. https://doi.org/10.1109/TSMC.2021.3113823
18. Huang R, Wang Y, Li Z, Lei Z, Xu Y (2022) RF-DCM: multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection. IEEE Trans Intell Transp Syst 23(1):630–640. https://doi.org/10.1109/TITS.2020.3017513

19. Fatima B, Shahid AR, Ziauddin S, Safi AA, Ramzan H (2020) Driver fatigue detection using Viola Jones and principal component analysis. Appl Artif Intell 34(6):456–483. https://doi.org/10.1080/08839514.2020.1723875
20. Wang C et al (2020) Spectral analysis of EEG during microsleep events annotated via driver monitoring system to characterize drowsiness. IEEE Trans Aerosp Electron Syst 56(2):1346–1356. https://doi.org/10.1109/TAES.2019.2933960
21. Zhu A, Cao S, Yao H, Jadliwala M, He J (2020) Can wearable devices facilitate a driver's brake response time in a classic car-following Task? IEEE Access 8:40081–40087. https://doi.org/10.1109/ACCESS.2020.2971632

Predicting the Heart Disease Using Machine Learning Techniques

Somya Goyal

Abstract Heart disease refers to the condition when the heart is not capable to push required amount of blood to the entire body. Heart disease (HD) is the prevailing reason behind deaths among the world-wide population. Early prediction of heart diseases can save lives. Predicting cardiovascular or heart disease in advance, a person can be warned beforehand, and the death can be prevented in turn. Machine learning (ML) has made a huge contribution to classify the population with heart disease from the healthy population. This paper proposes three heart disease prediction (HDP) models namely LOFS-ANN, LOFS-SVM, and LOFS-DT utilizing lion optimization-based feature selection (LOFS) method and three ML-based classifiers. The datasets used are from UCI repository. The comparative analysis reflects that the model LOFS-ANN performs best among all three models, with the values of 97.1% and 90.5% for AUC measure and accuracy measure, respectively. It can be concluded that the LOFS-ANN has a significant potential to predict heart disease after drawing its statistical comparison with the competing models.

Keywords Heart disease · Artificial neural network (ANN) · UCI Cleveland · Feature selection (FS) · Support vector machine (SVM) and area under the curve (AUC)

1 Introduction

Heart disease (HD) is the biggest reason behind the deaths all around the world. The WHO investigated into the statistics and reported that 17.7 deaths were caused due to cardiovascular diseases almost in 2015 throughout the world [1]. The early prediction of HD among population can be a potential help in saving lives by issuing warning and precautionary measures to the people. Machine learning (ML) techniques are playing a crucial role in heart diseases prediction (HDP) using the past collected patient data [2]. A wide range of ML techniques is available for developing the heart

S. Goyal (✉)
Manipal University Jaipur, Jaipur, Rajasthan 303007, India
e-mail: somyagoyal1988@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_21

disease predictors [3]. The patient datasets possess numerous attributes and not all worthy for predicting the heart disease. Feature selection (FS) facilitates to enhance prediction accuracy by removing the non-contributing and irrelevant attributes [4–8]. Bio-inspired algorithms are gaining popularity for the FS [9]. This study utilizes lion optimization (LO) algorithm originated from the social behavior of lion [10]. Lion optimization for feature selection (LOFS) has not yet been utilized in ML-based HDP domain. To carry out the research streamlined, following research goals are established-

R1To report the best ML-based HDP model among the proposed models to predict heart disease effectively.
R2 To establish the statistical validation of the work.

The paper is organized as follows—Sect. 2 discusses the literature related to this study. The experimental methods and setup are given in Sect. 3. The results of experiments are reported under Sect. 4. The research work is concluded under Sect. 5 bringing a light on the future work.

2 Literature Work

The survey on the work carried out in the literature of HDP applying the machine learning techniques has been summed up in this section. The survey is summarized as Table 1.

3 Research Methodology

The research methodology adopted for this work including the experimental methods and setup are briefed in this section.

This work utilizes three datasets from the UCI repository for experimental work [15]. The description to datasets attributes is given as under Table 2. The patient dataset is partitioned into training and testing datasets with 70–30 ratio. Then, lion optimization algorithm for feature selection (LOFS) [14] is applied to select the most significant features. The features selected using the LOFS algorithm for all three experimental datasets are listed as in Table 3. Then, the only selected features are fed to the ML-based classifiers for training purpose. The most renowned classification algorithms [2] are selected for the heart disease prediction (HDP) which are artificial neural network (ANN) [16], support vector machine (SVM), [17] and decision trees (DT) [18, 19]. Performance of all three proposed classifiers is recorded over all three datasets. Figure 1 depicts the proposed experimental model.

For the performance evaluation, ROC, AUC, and accuracy are considered [2, 3, 11–13, 16–21].

Table 1 Related work in the literature

S. No	Study	Dataset used	Technique used	Evaluation criteria	Inference drawn
1	Amin et al. [6]	UCI Cleveland dataset, UCI Statlog dataset	Decision tree, Naive Bayes, SVM, ANN	Accuracy	Improved performance of Naïve Bayes with logistic regression
2	Prakash et al. [7]	UCI Cleveland dataset	Optimal criteria for FS	Computational time, accuracy	Reduced execution time
3	Gokulnath et al. [8]	UCI Cleveland dataset	GA + SVM	ROC	Better performance achieved via GA-FS
4	Haq et al. [11]	UCI Cleveland dataset	7 ML + 3 FS	AUC, ROC, MCC	ML-based HDP has potential to assist doctors clinically
5	Bharti et al. [12]	UCI Cleveland dataset	Neural network	Accuracy, precision	Combined with deep learning to improve performance
6	Charles et al. [13]	UCI Cleveland dataset	ANN	Accuracy, precision	Improved performance via FS
7	Fitriyani et al. [14]	UCI Cleveland UCI Statlog	Clustering + SMOTE	Accuracy, Precision	Better performance achieved over classifiers

Table 2 Description of the datasets used

S. No	Dataset name	Number of features	Number of records
1	UCI Heart Disease Dataset (Cleveland) [15]	13	303
2	UCI Statlog (Heart) [19]	13	1024
3	UCI Heart Failure Clinical Dataset [20]	12	270

Table 3 Features Selected Using LOFS Algorithm

S. No	Dataset	Total features #	Selected features #	Features selected
1	UCI Heart Disease Dataset (Cleveland) [15]	13	8	2, 3, 7, 8, 9, 11, 12, 13
2	UCI Statlog (Heart) [19]	13	6	2, 3, 7, 9, 12, 13
3	UCI Heart Failure Clinical Dataset [20]	12	3	5, 8, 12

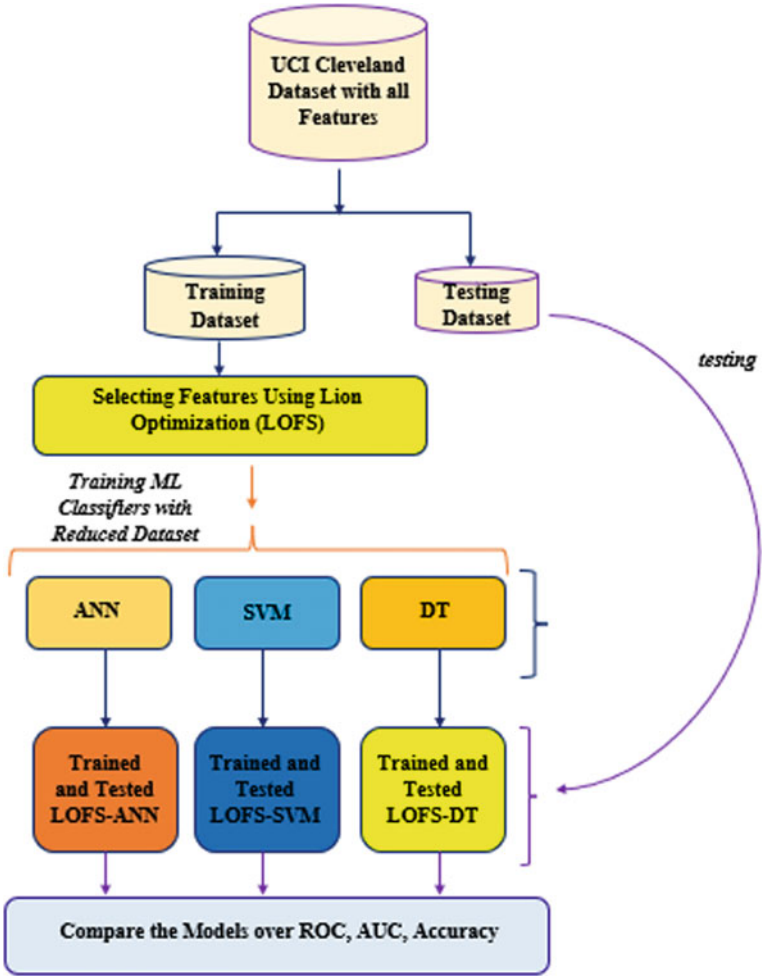


Fig. 1 Proposed heart disease prediction model with LOFS

4 Results and Discussion

This section reports the experimental results and the inferences drawn after analysis are listed out here.

Table 4 Comparison over AUC

S. No	Dataset	*LOFS-ANN*	*LOFS-SVM*	*LOFS-DT*
1	UCI Heart Disease (Cleveland)	*0.9413*	0.7459	0.8199
2	UCI Statlog (Heart)	*0.9714*	0.9286	0.9447
3	UCI Heart Failure Clinical	*0.924*	0.7113	0.8273

Table 5 Comparison over accuracy

S. No	Dataset	*LOFS-ANN*	*LOFS-SVM*	*LOFS-DT*
1	UCI Heart Disease (Cleveland)	*0.8911*	0.6106	0.7723
2	UCI Statlog (Heart)	*0.9005*	0.8859	0.878
3	UCI Heart Failure Clinical	*0.8595*	0.7023	0.7926

4.1 Finding the Best ML-Based HDP Model (R1)

A comparison is done among LOFS-ANN, LOFS-SVM, and LOFS-DT to find the best performer. First up, the AUC values are recorded over all three datasets for all the candidate models and reported as in Table 4. Next, the author records the accuracy measure (see Table 5). It is clear that LOFS-ANN performs best over accuracy criteria too. The results are plotted as Fig. 2 for visualization of comparative analysis.

To achieve the goal R1, ROC is considered for performance evaluation. The corresponding ROC plots for all three datasets—UCI Heart Disease Dataset (Cleveland) [15], UCI Statlog (Heart), and UCI Heart Failure Clinical Dataset are reported as Figs. 3, 4, and 5, respectively.

From the experimental results, it is seen that LOFS-ANN shows the best accuracy for predicting the heart disease in comparison with rest of the models.

Response to R1—The proposed LOFS-ANN performs best among the proposed models for all datasets.

4.2 Statistical Justification (R2)

To find the statistical proof, Friedman's test is conducted [20]. The result of test reflects upon whether the statistical proof for the goal R1 exists or not. The test is conducted with significance level of 5%. The results show that the value of p-statistic is less than 0.05 (see Fig. 6). Hence, it can be statistically validated that proposed LOFS-ANN-based HDP model is better than LOFS-SVM and LOFS-DT.

Response to R2—There exists statistical proof to validate the research work carried out in this paper.

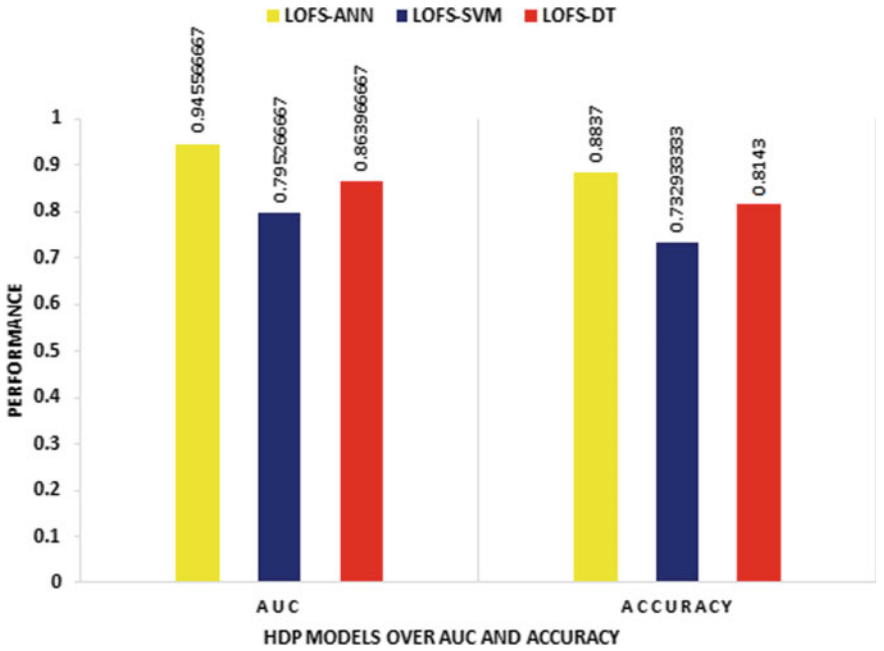


Fig. 2 Comparison of HDP models over AUC and accuracy

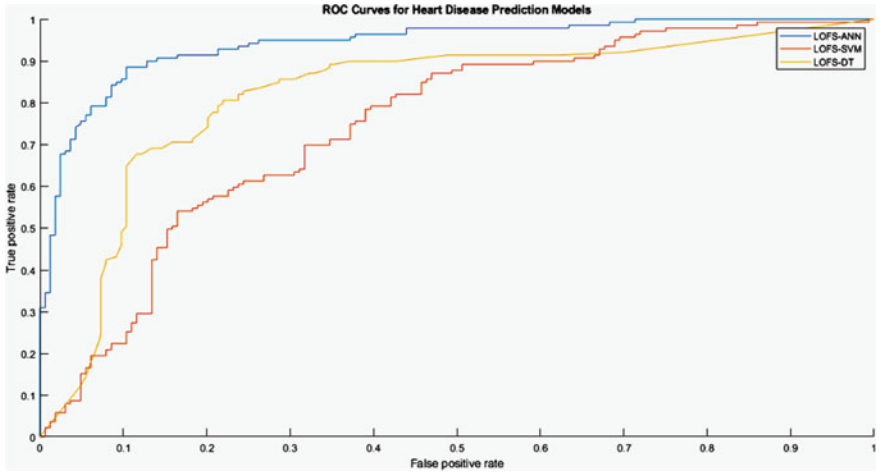


Fig. 3 ROC curve over UCI Heart Disease Dataset (Cleveland)

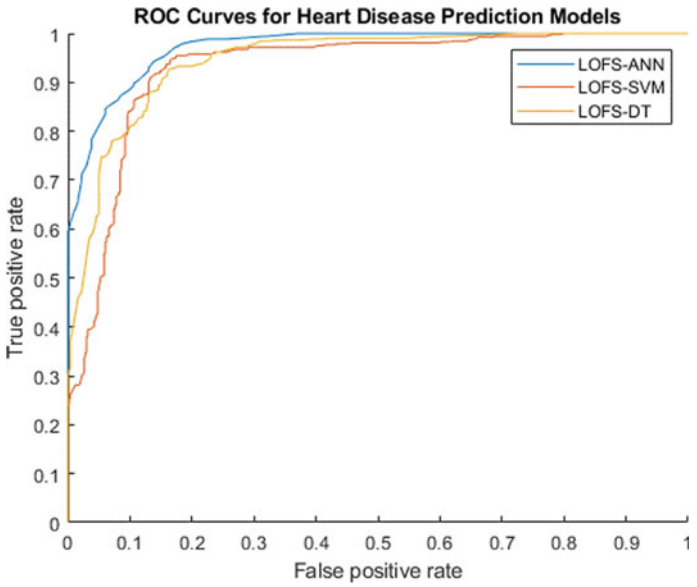


Fig. 4 ROC curve over UCI Statlog (Heart)

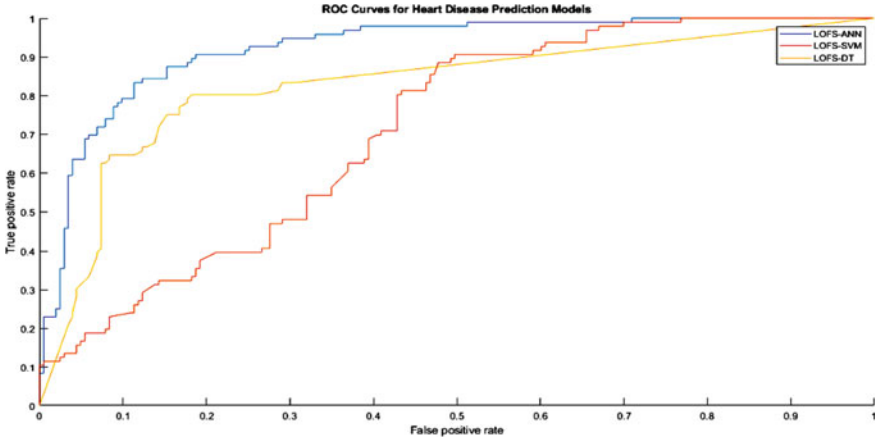


Fig. 5 ROC curve over UCI Heart Failure Clinical Dataset

Friedman's ANOVA Table

Source	SS	df	MS	Chi-sq	Prob>Chi-sq
Columns	6	2	3	6	0.0498
Error	0	4	0		
Total	6	8			

Fig. 6 p-statistic for Friedman test

5 Conclusion

Heart disease is the biggest reason of death in the entire world. If it is predicted well in advance and the patient is fore alarmed, then the lives can be saved. ML classification algorithms are being used for predicting the heart disease. The accuracy of the heart disease predictor is enhanced with the appropriate subset selection of the features from the total feature set—which are in good correlation with the target. In this paper, lion-based feature selection (LOFS) method has been utilized to select most significant features from three datasets—UCI Heart Disease Dataset (Cleveland), UCI Statlog (Heart), and UCI Heart Failure Clinical Dataset. These preprocessed data are fed for the training of three classifiers—ANN, SVM, and DT resulting into three HDP models-LOFS-ANN, LOFA-SVM, and LOFS-DT. The comparison is made among the performance of these proposed methods. The author concludes the work that the ANN with LOFS performs best for heart disease prediction.

Author proposes to replicate the work in the future with larger clinical datasets to contribute more accurate heart disease predictors for biomedical domain.

References

1. World Health Organization (WHO) (2017) Cardiovascular diseases (CVDs)—Key Facts
2. http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed 22 Mar 2022
3. Goyal S (2023) Software measurements with machine learning techniques-a review. Recent Adv Comput Sci Commun 16:1–17. https://dx.doi.org/10.2174/2666255815666220407101922
4. Safdar S, Zafar S, Zafar N et al (2018) Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. Artif Intell Rev 50:597–623. https://doi.org/10.1007/s10462-017-9552-8
5. Goyal S (2022) FOFS: firefly optimization for feature selection to predict fault-prone software modules. In: Nanda P, Verma VK, Srivastava S, Gupta RK, Mazumdar AP (eds) Data engineering for smart systems. Lecture Notes in Networks and Systems, vol 238. Springer, Singapore. https://doi.org/10.1007/978-981-16-2641-8_46
6. Amin MS, Chiam YK, Varathan KD (2019) Identification of significant features and data mining techniques in predicting heart disease. Telemat Inform 36:82–93. https://doi.org/10.1016/j.tele.2018.11.007
7. Prakash S, Sangeetha K, Ramkumar N (2019) An optimal criterion feature selection method for prediction and effective analysis of heart disease. Cluster Comput 22(s5):11957–11963. https://doi.org/10.1007/s10586-017-1530-z
8. Gokulnath CB, Shantharajah SP (2019) An optimized feature selection based on genetic approach and support vector machine for heart disease. Cluster Comput 22(s6):14777–14787. https://doi.org/10.1007/s10586-018-2416-4
9. Darwish A (2018) Bio-inspired computing: algorithms review, deep analysis, and the scope of applications. Future Comput Inform J 3(2):231–246, ISSN 2314-7288. https://doi.org/10.1016/j.fcij.2018.06.001
10. Yazdani M, Jolai F (2016) Lion optimization algorithm (LOA): a nature-inspired metaheuristic algorithm. J Comput Design Eng 3(1):24–36, ISSN 2288-4300. https://doi.org/10.1016/j.jcde.2015.06.003

11. Haq AU, Li JP, Memon MH, Nazir S, Sun R (2018) A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mobile Inform Syst
12. Bharti R, Khamparia A, Shabaz M, Dhiman G, Pande S, Singh P (2021) Prediction of heart disease using a combination of machine learning and deep learning. Comput Intell Neurosci
13. Benhar Charles V, Surendran D, SureshKumar A (2022) Heart disease data based privacy preservation using enhanced ElGamal and ResNet classifier. Biomed Signal Process Control 71(Part B):103185, ISSN 1746-8094. https://doi.org/10.1016/j.bspc.2021.103185
14. Fitriyani NL, Syafrudin M, Alfian G, Rhee J (2020) HDPM: an effective heart disease prediction model for a clinical decision support system. IEEE Access 8:133034–133050. https://doi.org/10.1109/ACCESS.2020.3010511
15. Goyal S (2022) Genetic evolution-based feature selection for software defect prediction using SVMs. J Circuits Syst Comput 31(11):2250161. https://doi.org/10.1142/S0218126622501614
16. Goyal S (2022) 3PcGE: 3-parent child-based genetic evolution for software defect prediction. Innovations Syst Softw Eng. https://doi.org/10.1007/s11334-021-00427-1
17. UCI Machine Learning Repository: Heart Disease Data Set.: Archive.ics.uci.edu. http://archive.ics.uci.edu/ml/datasets/Heart?
18. Goyal S (2021) Effective software defect prediction using support vector machines (SVMs). Int J Syst Assur Eng Manag. https://doi.org/10.1007/s13198-021-01326-1
19. Goyal S (2022) Static code metrics-based deep learning architecture for software fault prediction. Soft Comput pp 1–33. https://doi.org/10.1007/s00500-022-07365-5
20. Goyal S (2021) Predicting the defects using stacked ensemble learner with filtered dataset. Autom Softw Eng 28:14. https://doi.org/10.1007/s10515-021-00285-y
21. Goyal S (2021) Handling class-imbalance with KNN (neighbourhood) under-sampling for software defect prediction. Artif Intell Rev. https://doi.org/10.1007/s10462-021-10044-w

Synthetically Generated High-Resolution Reflected Ultra Violet Imaging System (RUVIS) Database Using Generative Adversarial Network

Ritika Dhaneshwar, Mandeep Kaur, and Manvjeet Kaur

Abstract In this era of deep learning and artificial intelligence, where a large scale high-quality dataset is a must, we present a high resolution (1024*1024) synthetically generated latent fingerprint RUVIS dataset with pore level details. It will assist the research community to overcome the privacy issue concerns which hinders the large scale manual collection of fingerprint datasets. For this, the training dataset has been collected using RUVIS equipment which works on the principle of reflection of ultraviolet light, whereas synthetic prints are generated using StyleGan2-Ada. Our experimentation suggests that the synthetically generated latent fingerprints are not only diverse and unique but also highlights pore level details. Further, the experimental analysis has been done using structural similarity index (SSIM) and multi-scale structural similarity (MS-SSIM).

Keywords Latent fingerprint · Enhancement · Reconstruction · Segmentation · Matching · Generative adversarial network

1 Introduction

Fingerprints are amongst the most widely used evidence for recognition and authentication of individual identity. Further, law enforcement agencies consider it be cardinal evidence for convicting criminals. Depending on the nature of the evidence surface, there are various methods for uplifting a fingerprint-like powder method, sodium fuming method, etc. [1]. After the upliftment of evidence, we perform various preprocessing operations like enhancement, reconstruction, segmentation, etc., to improve

R. Dhaneshwar (✉) · M. Kaur
University Institute of Engineering and Technology (UIET), Panjab University, Chandigarh, India
e-mail: ritikadhaneshwar@gmail.com

M. Kaur
e-mail: mandeep@pu.ac.in

M. Kaur
Cyber Security Research Centre (CSRC), Punjab Engineering College (PEC), Chandigarh, India
e-mail: manvjeet@pec.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_22

its quality which in turn helps in accurate matching. Fingerprint evidence can be collected either in a controlled environment or an uncontrolled environment. Collection in a controlled environment is done using high-quality scanners with minimal distortions, whereas in real-world situations, the evidence is of poor quality because of the uncontrolled environment. Latent fingerprints are unintentional or accidental prints that are left at crime scenes. They are generally formed due to natural secretions of the human body like traces of sweat, oil, etc. Due to their poor quality, it requires a lot of preprocessing where machine learning could play an important role. So to improve the performance of latent fingerprint recognition systems, researchers are moving towards applying deep learning approaches. These machine learning approaches require a large quality dataset for training their models, that is minimally available publically. This challenge before the research community is discouraging them to apply emerging machine learning approaches that could otherwise help to immensely improve the performance parameters of their approach.

We propose a novel approach that would help us in generating realistic looking and high-resolution latent fingerprints. These synthetic fingerprints will overcome the legal issues of privacy that are associated with the real biometric dataset. Our goal is to encourage the use of deep learning frameworks and approaches in this area of research. This would help in improving fingerprint recognition reliability and efficiency. Our contributions are to use the StyleGan2-Ada augmentation feature for generating high-resolution fingerprints (pore level details) with a limited training dataset. The visual analysis shows that results are realistic with comparable quality. In Sect. 2, we present the literature review of the existing synthetic fingerprint datasets generation approaches. Section 3 deals with the approach adopted for generating fingerprints with details regarding training and augmentations. Experimental results are discussed in Sect. 4 of the literature.

2 Related Work

A novel approach was proposed by Cao et al. [2] in which 512 × 512 rolled fingerprint images are generated using the generative adversarial network (GAN). The improved-WGAN (I-WGAN) architecture utilizes Wasserstein distance for calculating the value function which leads to better performance than the available generative approaches [2]. A generative adversarial network based on a machine learning framework was proposed in 2018 [3]. By this proposition, they were trying to capture complex texture representations of fingerprint images to ensure the generation of realistic-looking images which helps in overcoming the major drawback of classical statistical models.

Further, a new approach was proposed which was based on generative adversarial networks and the super resolution methods [4]. This approach aims to generate large scale high-fidelity synthetic fingerprints. The probability distribution of fingerprints is estimated using generative adversarial networks, whereas to ensure fine-grained textures super resolution method is used. The approach proposed by Fahim et al.

[5] is based on the generative adversarial network where a combination of residual network and spectral normalization is used. Spectral normalization helps to ensure the stability of the network while training the model. Further, the issue of vanishing gradient is handled efficiently using the proposed average residual network [5]. This approach utilizes spectral bounding in the input and the fully connected layers.

A Clarkson fingerprint generator (CFG) was proposed in 2021, which helps in synthesizing full plain impressions of 512 × 512 pixels [6]. It is based on progressive growth-based generative adversarial networks (GANs). Using this technique, they can generate high-fidelity realistic fingerprints which are diverse and resemble the training dataset in terms of quality. The author proposes a PrintsGAN model which helps in the generation of realistic-looking synthetic fingerprints [7]. This model comprises of series of steps that lead to the generation of high-quality rolled fingerprints. The first step in this approach is to generate a binary Master-Print IID using a random noise vector. For the learning of mapping to binary prints, this approach utilizes BigGAN architecture. This generator aims to generate synthetic binary fingerprints.

3 Synthetic Latent Fingerprint Generation

Generative adversarial networks (GANs) [9] are deep learning-based generative models where two competing neural networks that are generator and discriminator compete with each other to generate synthetic data. The generator generates random output and proposes it to be real, whereas a discriminator detects and label the data as real/fake. In the scenario, where the data is classified as fake, the weights of the network are updated using the back-propagation approach. There are various flavours of GAN's CycleGAN [10], deep convolutional GAN's(DCGAN) [11], conditional GAN(CGAN) [12], semi-supervised GANS(SGAN) [13], etc., with wide scale of applications like cloning [14], synthetic data generation [15], handwriting generation [16] and various other audio and speech applications [17]. Though with such a wider application base, some of its drawbacks are—GANs are very resource hungry. It is required to be fed with tons of data to get quality results. Further, to process such a huge amount of data, we require high-end resources, which is practically not possible for every researcher. Moreover, most of the GAN's do not generate high-resolution images, which is the need of the hour.

3.1 Adaptive Discriminator Augmentation (StyleGan2-Ada)

StyleGan2-Ada is a network architecture (Fig. 1) with the ability to train with a limited dataset without facing the problem of discriminator overfitting [8]. It provides us with better resolution results as well as more realistic images. They work on the principle

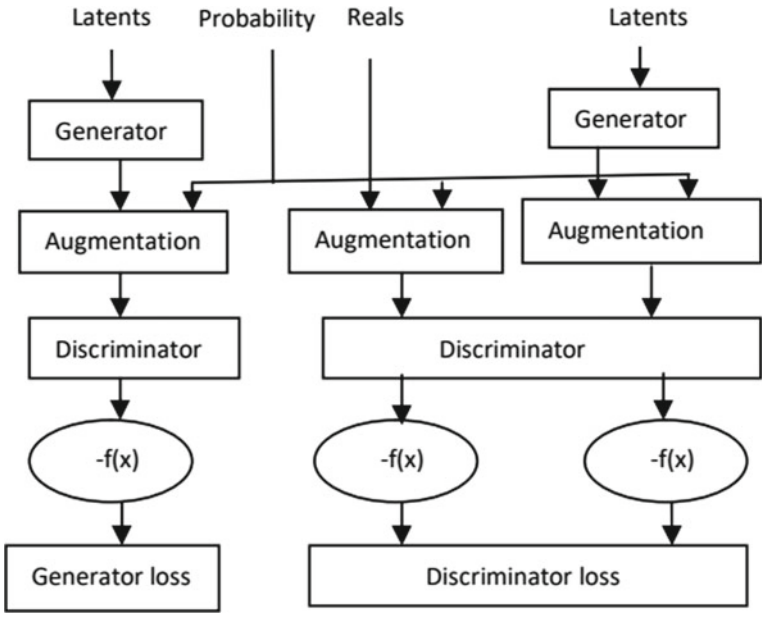


Fig.1 Block diagram of StyleGan2-ADA [8]

of latent space and style network. Style network helps to manipulate image style which leads to the generation of the more realistic texture image.

The basic working principle of StyleGan2-ADA is depicted in Fig. 1. The loss function used is non-saturating logistic loss [8] $f(x) = \log(\text{sigmoid}(x))$. Further, a diverse set of augmentations is applied on images which are directed by augmentation probability p.

3.2 StyleGan2-Ada Training Dataset

The training dataset samples as depicted in Fig. 2 are collected using an optical touchless device called RUVIS [18]. This works on the principle of reflection of ultraviolet rays. It helps in searching, identifying, developing and preserving the latent fingerprints. The advantage of using this technology is that it captures the evidence without touching it, which is not the case with other available approaches for fingerprint upliftment. Due to its high-resolution capturing, the training set includes samples with level-3 features as shown in Fig. 3 (pore level details).

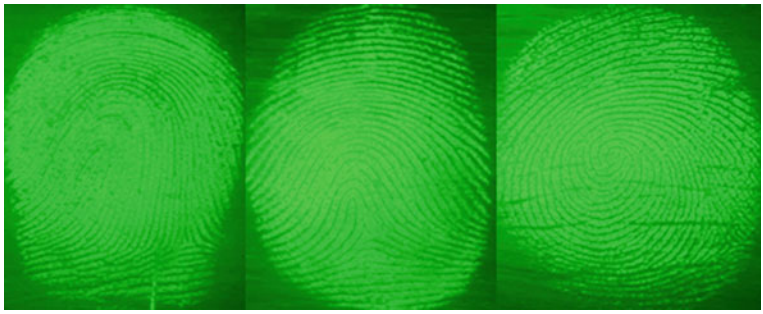


Fig. 2 Sample training dataset images (1024*1024)

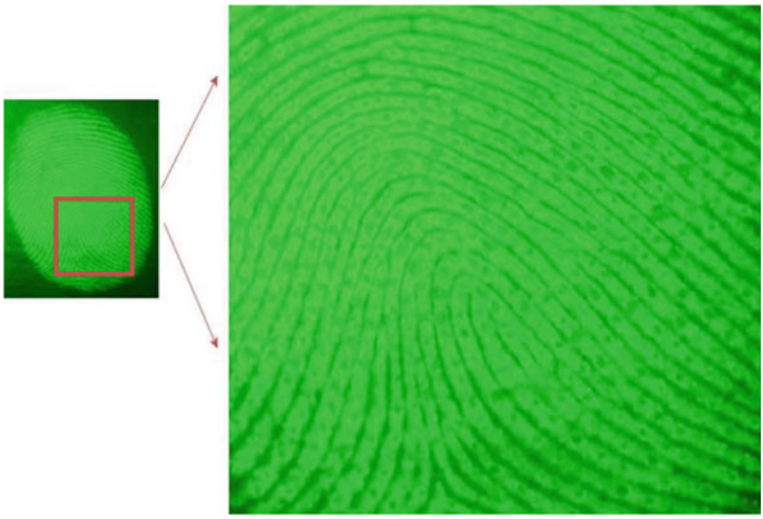


Fig. 3 Sample training dataset with L3 features

3.3 Adaptive Discriminator Augmentation

While training with small datasets (~1−10 k images), the discriminator may come across similar images repetitively, which leads to the discriminator memorizing real images. This may hamper the generators ability to create new images. So, to deal with this problem of overfitting StyleGan2-Ada provides us with the solution, i.e. augmentation of images that are fed to the discriminator. Augmentations have been performed to create additional imagery or additional data. The augmentation performed by us during the training of our model is 'bg' as per StyleGan2-Ada specifications [8]. It helps us to apply the following augmentations like Xflip, Xint, scaling, rotation of images, etc. Apart from these augmentations, we have performed a mirror operation,

which flips our images from left to right as well as top to bottom. These augmentation and mirror operations are fully automated without any manual intervention. Figure 4 depicts some of the samples with augmented data. Further, ADA also helps us in adjusting the probability of augmentations in our dataset. This probability is automatically adjusted based on the degree of overfitting as the training proceeds.

To maintain the high resolution of our dataset as well as to save time and resources, we utilized the Flickr-Faces-HQ FFHQ1024 [19] model for transfer learning as shown in Fig. 5. This technique helps us to generate synthetic images with a resolution of 1024*1024, due to which we can extract pore level details. The biggest advantage of using this approach is to train our model with a small dataset, without any compromise with the quality of an image.

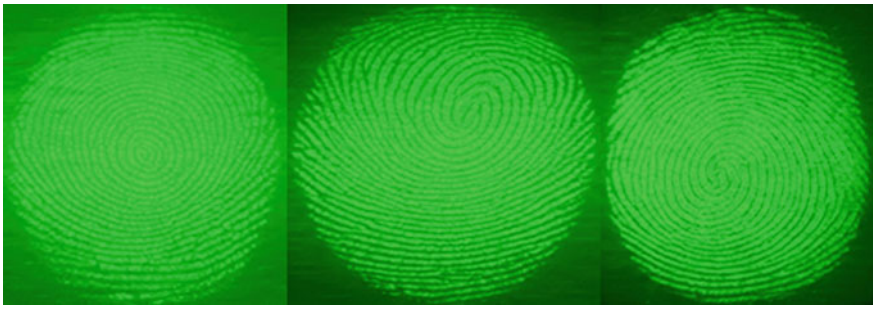


Fig. 4 Samples of augmented latent fingerprints

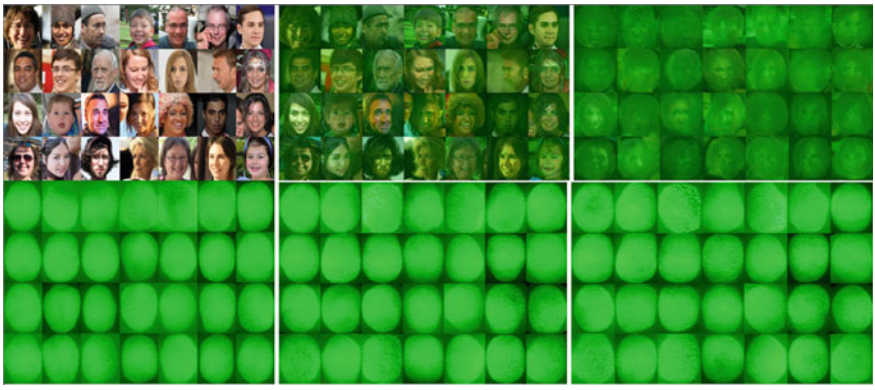


Fig. 5 Subsequent outputs obtained during transfer learning

4 Experimental Results and Discussions

In this section, we would evaluate our experimental results based on two verticals. First of all, we would present the qualitative analysis of our work which would be further quantified using multi-scale structural similarity [20] and structural similarity index [21].

4.1 Visual Analysis

To verify the quality of generated latent fingerprints, we did a visual inspection of our samples. This helped us to evaluate our generated samples against the input training dataset. Through this analysis, we were able to conclude that the generated samples are of high resolution with level-3 (L3) features like pores. These L3 traits increase the reliability and efficiency of fingerprint recognition [22]. Figure 6 depicts the visual comparison with L3 traits. Hence, the quality of generated images is comparable to the input dataset visually.

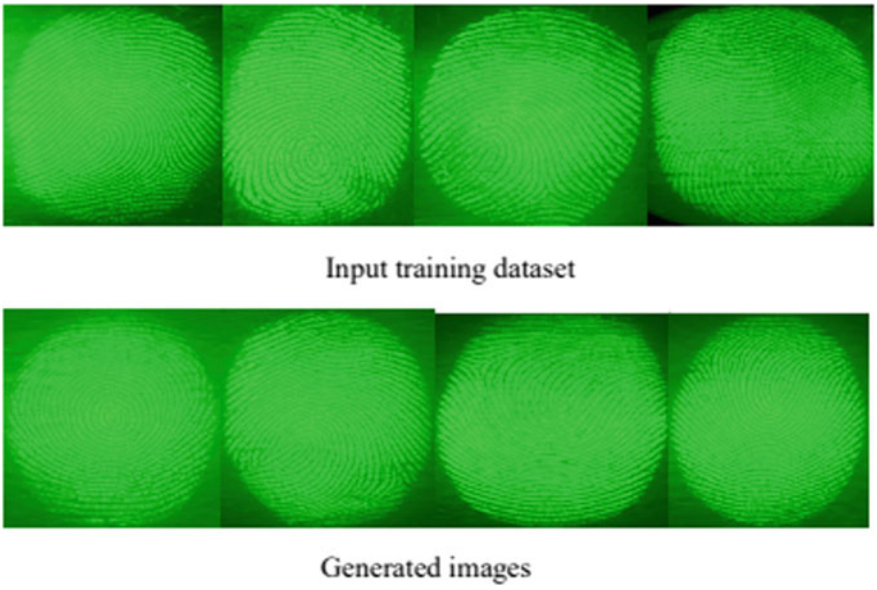


Fig. 6 Outputs generated during training of StyleGan2-Ada

Table 1 MS-SSIM and average SSIM score comparison

Score	Range
MS-SSIM score	0.28–0.35
Average SSIM score	0.26–0.30

4.2 Multi-Scale Structural Similarity

There are many full reference image quality parameters like peak signal to noise ratio, root means square error, global relative error, etc., used by researchers for experimental assessment. However, these indices are not that accurate when applied to latent fingerprints, because they do not capture the intricate details that are imperative while evaluation. So, to get an accurate evaluation of our results, we computed multi-scale structural similarity. This is based on the assumption that the human visual system is adapted for retrieving structural information from an image, which gives accurate results for the perceived quality of an image. The contrast comparison, structure comparison and luminance comparison are the three main components using which a similarity score is calculated. This method takes into account the texture level details of a sample which further improves its reliability of diversity analysis. This metric returns a score between 0.0 and 1.0. The higher the score per batch, the lower the amount of diversity amongst the generated images [23].

Table 1 depicts the average MS-SSIM and SSIM scores which are calculated between the original training images and synthetic fingerprints. Corresponding to one generated image, there are 1000 scores which are then averaged to get a quantifiable value. A similar procedure is adopted for the rest of the fake images. The best MS-SSIM score achieved by our study is 0.16. Since our mean SSIM score of synthetic images did not exceed our MS-SSIM scores, hence, justifies our results.

5 Conclusion

Through this research, we demonstrate that StyleGan2-Ada is amongst the most suitable approach for the generation of synthetic data with a limited dataset. It helps overcome the issue of a large training dataset (otherwise overfitting) which is a drawback amongst most of the available GAN variants. Using this approach, we can generate high-resolution synthetic fingerprints with pore level details which would encourage the research community to further explore deep learning approaches and frameworks in this domain.

Acknowledgements This research is supported by the RUVIS dataset of Cyber Security Research Centre, Punjab Engineering College, Chandigarh, India.

References

1. Dhaneshwar R, Kaur M, Kaur M (2021) An investigation of latent fingerprinting techniques. Egypt J Forensic Sci 11(1):1–15
2. Cao K, Jain A (2018) Fingerprint synthesis: evaluating fingerprint search at scale. In: International conference on biometrics (ICB). IEEE, pp 31–38
3. Minaee S, Abdolrashidi A (2018) Finger-GAN: generating realistic fingerprint images using connectivity imposed GAN. arXiv preprint arXiv:1812.10482
4. Riazi MS, Chavoshian SM, Koushanfar F (2020) SynFi: automatic synthetic fingerprint generation. arXiv preprint arXiv:2002.08900
5. Fahim MANI, Jung HY (2020) A lightweight GAN network for large scale fingerprint generation. IEEE Access 8:92918–92928
6. Bahmani K, Plesh R, Johnson P, Schuckers S, Swyka T (2021) High fidelity fingerprint generation: quality, uniqueness, and privacy. arXiv preprint arXiv:2105.10403
7. Cao K, Jain A (2018) Fingerprint synthesis: evaluating fingerprint search at scale. In: 2018 International conference on biometrics (ICB). IEEE, pp 31–38
8. Karras T, Aittala M, Hellsten J, Laine S, Lehtinen J, Aila T (2020) Training generative adversarial networks with limited data. arXiv preprint arXiv:2006.06676
9. Gonog L, Zhou LY (2019) A review: generative adversarial networks. In: 14th IEEE conference on industrial electronics and applications (ICIEA). IEEE, pp 505–510
10. Zhu JY, Taesung P, Phillip I, Alexei AE (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp 2223–2232
11. Radford A, Luke M, Soumith C (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434
12. Mirza M, Osindero S (2014) Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784
13. Odena A (2016) Semi-supervised learning with generative adversarial networks. arXiv preprint arXiv:1606.01583
14. Terziyan V, Gavriushenko M, Girka A, Gontarenko A, Kaikova O (2021) Cloning and training collective intelligence with generative adversarial networks. IET Collaborative Intell Manuf
15. Vega-Márquez B, Cristina RE, José CR, Isabel NC (2019) Creation of synthetic data with conditional generative adversarial networks. In: International workshop on soft computing models in industrial and environmental applications. Springer, Cham, pp 231–240
16. Ghosh A, Biswarup B, Somnath BRC (2017) Handwriting profiling using generative adversarial networks. In: Proceedings of the AAAI conference on artificial intelligence, vol 31, no 1
17. Donahue C, Bo L, Rohit P (2018) Exploring speech enhancement with generative adversarial networks for robust speech recognition. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 5024–5028
18. Singla N, Kaur M, Sofat S (2020) Latent fingerprint database using reflected ultra violet imaging system. Procedia Comput Sci 167:942–951. https://doi.org/10.1016/j.procs.2020.03.393
19. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 4401–4410
20. Wang Z, Eero PS, Alan CB (2003) Multiscale structural similarity for image quality assessment. In: The Thrity-Seventh Asilomar conference on signals, systems & computers, 2003, vol 2. IEEE, pp 1398–1402
21. Brune D, Edward RV, Zhou W (2011) On the mathematical properties of the structural similarity index. IEEE Trans Image Process 21(4):1488–1499
22. Agarwal D, Bansal A (2021) A utility of pores as level 3 features in latent fingerprint identification. Multimedia Tools Appl 2021:1–20
23. Fahi M, Ho YJ (2020) A lightweight GAN network for large scale fingerprint generation. IEEE Access 8(2020):92918–92928

Understanding and Comparative Analysis of Consensus Algorithms

Neha A. Samsir and Arpit A. Jain

Abstract Blockchain technology has emerged as a new security technology that offers a variety of benefits to numerous businesses. Blockchain is working on the concept of distributed ledger technology. Multiple industries are adopting blockchain technology due to its different features like distributed networks, consensus mechanisms, and transparency. Blockchain technology gives a great impact on the business environment. Blockchain technology does not involve third-party interference and provides security and transparency to data. The data inside the blockchain is immutable, and no one will be modified it or delete it. Smart contracts and consensus algorithms are the major components of blockchain technology. A smart contract is a sort of computer software that establishes agreements between the nodes of a blockchain network's participants. Few major consensus algorithms are PoW, PoS, PBFT, DPoS, and tendermint. This paper evaluates the different consensus algorithms and principles behind using it. Paper demonstrates the steps to be followed for the implementation of consensus algorithms.

Keywords Blockchain · Security · Mining · Hashing · Consensus algorithms

1 Introduction

The blockchain was first introduced as a Bitcoin, a type of cryptocurrency created by "Satoshi Nakamoto" in 2008. Blockchain operates on a peer-to-peer basis. The traditional payment system works based on a centralized system, and the entire transactions run through the central authority. It may cause problems with security, efficiency, and transaction cost. A decentralized network can solve these problems and gives equal authority to all nodes in the network. All nodes achieve their goals based on consensus algorithms and smart contracts [1]. A blockchain, which is a

N. A. Samsir (✉) · A. A. Jain
GLS University, Gujarat, India
e-mail: nehasamsir@gmail.com

A. A. Jain
e-mail: arpit.jain@glsuniversity.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_23

type of distributed ledger, saves all of the ledger's transactions in a collection of blocks. Blockchain design has overcome the intermediaries and vulnerable security attacks issue. Blockchain has basic three features which are decentralization, transparency, and immutability. Decentralization means transferring the decision-making and control from a central authority to distributed nodes. All nodes in a network take participants in all transactions, and no single authority is responsible for any activity. In the blockchain, transparency word indicates that anyone can join the network, and all nodes of the network access the data. The transparency feature makes data available to everyone at any time, and immutability is about data which means the data inserted in the blockchain is unaltered and unchangeable. Every transaction in blockchain uses a cryptographic principle that makes data unchangeable. Blockchain is one type of technology that includes many techniques which include mathematic algorithms, cryptography, incorporating peer-to-peer networks, and different consensus algorithms. Generally speaking, blockchain is divided into three categories: public, private, and consortium. A public blockchain, also known as a permissionless blockchain, uses a proof of work consensus method to operate. All nodes on a public blockchain have public access to the transaction or source code. A permission blockchain, also known as a private blockchain, is used to approve a system and may only be accessed by those who have been granted access. Consortium blockchain includes many institutions where the authentication of transactions is verified through only a particular organization. This type of blockchain is used for the financial and industrial sectors [2, 3]. Consensus is a fundamental concept in blockchain technology that ensures a tamper-proof environment for the network and only one version of agreement among all nodes. The attacker will find it difficult to detect a manipulated block in the blockchain since all nodes in the network must reach an agreement based on decentralization [4, 5]

2 Literature Review

Blockchain is a very vase technology. Smart contracts and consensus algorithms play a major role in becoming technology more impactful in different application areas. A huge amount of literature reviews are available on blockchain smart contracts and consensus algorithms. The advantage of blockchain has been seen in various sectors which are as follows.

Jalal et al. [6] give the basics of consensus algorithms. They explain the process of choosing the best consensus protocol for the system. They also discussed the classification of consensus algorithms. Proof of contribution (PoC), Ripple, Stellar consensus protocol (SCP), proof of stack velocity (PoSV), proof of stake Casper (PoS Casper), proof of participation and fee (PoPF), cloud PoS, delegated PoS, and proof of disease are some of the consensus methods which are covered.

Hasselgren et al. [7] discussed the fundamental concept of blockchain, as well as its major attributes, blockchain kinds, consensus methods, and smart contracts. The writers of this study concentrated on blockchain's application in the healthcare

sector. Data Provence and data integrity is the basic key component of blockchain technology. The authors of this work cover several considerations while assessing or constructing the system's architecture, including search strategy, selection process, data extraction, data analysis, data extraction, and quality assessment. Following the research, the authors discuss various technical considerations for developing a blockchain, including blockchain type, a blockchain platform, consensus algorithm, and smart contracts.

Chaudry and Yousaf [8] discussed the basic introduction of blockchain technology and cryptocurrency. They give an overview of the Bitcoin and Ethereum blockchain. The authors gave an overview of distributed systems consensus and blockchain. They list crash failures, transitory failures, omission failures, byzantine 3 failures, software failures, temporal failures, and environmental perturbations as examples of failures. They also compare consensus algorithms and classify them based on blockchain type, scalability, adversary tolerance model, performance-related factors, communication model and complexity, energy consumption, mining and consensus category, and finality of consensus.

Mingxiao et al. [9] discussed the basics of consensus algorithms and provide an understanding of different algorithms which are PoW, PoS, DPoS, PBFT, and RAFT. They analyze different consensus algorithms based on their performance. Also, mention some of the drawbacks of PoW algorithms, such as resource waste, transaction speed, and hashing power concentration. Blockchain is separated into three categories in terms of application: public blockchain, private blockchain, and permissioned blockchain. Finally, the authors suggest that blockchain technology has the following characteristics: security, stability, decentralization, and non-modifiability.

3 Blockchain Technology

3.1 Working of Blockchain

Blockchain is a huge technology that contains many phases. Blockchain transactions or data are stored in the format of blocks. Each block is connected to the other by creating a chain of blocks. All blocks in the blockchain are stored in a chronological model. Blockchain technology is one type of transactional database, and it can be created by various computers. In a blockchain network, each resource is represented by a node. Before blockchain, we need a central authority or intervention mechanism to handle all the transactions. This technology gives users authority to control their information and maintain the information regarding their transactions. A Blockchain network removes the central host or authority. Due to no central authority, single failure does not harm the system. It increases the system's reliability, durability, and longevity. All users can view their transactions with blockchain technology, and no one can alter them. The method of storing transactions in the blockchain is depicted in Table 1. Each block in the blockchain is validated and chained separately.

Table 1 Working of blockchain

Fields	Description
Transaction definition	This is the first step of blockchain. Initially, the transaction is created by the sender, then the sender also has the receiver's public address information and cryptographic digital signature. A digital signature verifies the creditability and validity of a transaction
Transaction authentication	The transaction authentication phase is held temporarily because if the nodes validate the transaction it is used to create a block as a node. Nodes perform the message validation process by decrypting digital signatures with the use of cryptography techniques
Block creation	In the block creation stage, it checks the pending transaction first. If a pending transaction is used by another node in the network, then it will be updated on the block or ledger. The updated block is shown to the nodes. This block is shown to another node
Block validation	The block validation process is used to validate the block by particular acquiescence. If a node receives a request for update block validation, then they perform a repetitive process node seeks another node's acquiescence in the network to verify the block
Block chaining	After the validation process, the validated block is added to the present blocks and created a chain of the blocks that are called the blockchain

3.2 Consensus Algorithms

To overcome two difficulties, double-spending and byzantine general problems, a blockchain application is required. A consensus algorithm is a kind of distributed system or process agreement that is based on a single data value. Blockchain is a decentralized network, and it aims to give security and immutability to data. All parties in the blockchain are working based on a common agreement that is present in the same ledger. Through the consensus algorithm, all parties in the blockchain come on to a common agreement on the present data in the ledger. Consensus algorithms are essential because they maintain data integrity and security of data. A few common consensus algorithms which are used in blockchain are as follows [9].

Proof of Work (PoW)
The Bitcoin network uses a pure consensus process known as proof of work. Integration with the Bitcoin PoW protocol applies to miners and miners solving a computationally hard puzzle. A PoW block header contains solving a computationally hard puzzle. A PoW block header contains the hash value, and it will be calculated for each node. The nonce value is contained in the block header, and the miner adjusts the nonce value regularly to obtain the various hash values. A nonce is a form of an abbreviated number that is added to a hashed or encrypted block on the blockchain and is only used once. A number that is greater than or equal to a hashed block header is used as the target value. The target hash is used to establish the input difficulty level, and it can be adjusted to guarantee that blocks in the blockchain are processed

efficiently. When one node reaches a specified target value, it is broadcast to the other nodes, and a hash value is mutually acknowledged by all nodes in the proof of work idea. Miners append the new block to the blockchain if the block is validated. Miners are identified by hash values, and PoW is known as mining.

PoW follows the following algorithm
1. Start
2. User request to the server to add a new block in a blockchain
3. Server generates the mathematical puzzle that is known as a hash number
4. If hash number < target value Proof of work achieved Else Increase the nonce value
5. If proof of work achieved A new block added to the blockchain Else Block is canceled
6. Block is added, then rewards are given to the miner
7. Stop

Proof of Stack (PoS)

A selection of approving nodes is known as proof of stack, and it is determined by the stack of each node in the blockchain. It is a more energy-efficient version of PoW. Validators are chosen as miners in proof of stake, and the mining process is known as minting or forging. When compared to proof of work, it is more decentralized. In cryptocurrencies, the stack represents the balance of a certain currency. Many blockchains started with PoW and are now switching to PoS. In digital money, coinage refers to the value multiplied by the amount of time since it was created. The longer the currency, the more network rights a node can gain. Holders of coins will receive a reward based on their coinage. In this algorithm, mining needs accounting rights. It is derived from this formula:

$$\text{proof hash} \ < \ \text{coinage} \ * \ \text{target}$$

The hash value of the weight factor, which is a mix of the fuzzy sum of the current time and the unspent output value, is contained in the proof hash. The limit is determined by each node's hashing power. The PoS algorithm encourages coin holders to keep their coins for longer periods. PoS is superior to PoW due to the concept of coinage. If attackers wish to assault the blockchain, they must first gather a huge quantity of currencies and keep them for a long period. The main benefit of using PoS is that it raises the attack difficulty.

PoS follows the following algorithm

1. Start
2. Validator stakes some amount of money on a particular transaction
3. According to their stakes validator selected
4. Validator creates the block for transaction
5. If a block is valid
   Block is added to the blockchain
   Else
   Reward and stakes amount not given to validator
6. Validator receives the reward
7. Stop

Practical Byzantine Fault Tolerance (PBFT)

Byzantine fault tolerance is a useful way to deal with transmission faults [1]. Hyperledger fabric has used a byzantine fault tolerance consensus algorithm for handling malicious attacks. In PBFT, a new block is determined in one round, and the primary is chosen in each round based on particular rules. We can use it to complete a transaction in the form of an order. Each procedure in PBFT is broken down into three phases: pre-prepared, prepared, and commit. All nodes must be known to the network in PBFT, which is used to restrict the use of the consensus process in a public blockchain. In each step, two-thirds of the network's nodes are required [6]. PBFT was created to reduce complexity to a polynomial level, resulting in increased efficiency. The five stages of the PBFT process are request, pre-prepare, prepare, commit, and reply.

PBFT follows the following algorithm

1. Start
2. All available nodes in a network are sequentially ordered
3. One node is set as a master server node in a round-robin manner
4. Client request to the master server node
5. Master server node invokes the service operation
6. Master server node broadcasts the replicas to the other nodes
7. Replicas execute the request
8. If the majority of nodes accept the request
   It will enter to commit the state
   Else
   Invalid the block
9. Each node in the commit state sends the commit message to other nodes
10. Master server node replies to the client for the success
11. Stop

Delegated Proof of Stack (DPoS)

During the design phase of Bitcoin, all nodes or participants mine using the CPU. Each node has the same hashing power as the others in the network and can participate in the blockchain's decision-making. Bitcoin is getting popularity and getting appreciation 6 from the users. With the increasing popularity of blockchain, the mining

concept was invented. A large number of mining machines divided the hashing power into groups of participants. The ordinary miners do not get the opportunity always to mine the block. Bitshare is using the DPoS algorithm. In the DPoS algorithm based on stake, each node selects witnesses. According to accounting rights, the highest number of witnesses in the DPoS network participated in the campaign and received the most votes. In DPoS consensus, the user gives a vote to another user to give power to another entity to give a vote on the behalf of another one. Selecting a witness is responsible for creating a new block and verifying the transaction. If the selecting witness is verified and signs the transaction, then they receive rewards. If they failed to verify and sign the transaction, then no rewards are assigned to the witness. The reward is added up to the next witness who verifies the block and this block is known as the stolen block. In comparison with PoS and PoW, DPoS is a more efficient and power-saving consensus [5].

DPoS follows the following algorithm

1. **Start**
2. voting for electing witness panel (N^i number of votes)
3. if N^i = (hash (previous block header), nonce)
   Hash (hash (previous block header), nonce) > D (where D is the value of difficulty)
   Else
   Increase the nonce value
4. System broadcast the consensus nodes N^c
5. **Tokens allocate as a vote for witnesses**
6. **Perform quick sort**
7. $N^w \leftarrow N^c$ (where N^w is a set of witness nodes)
8. $N^A \leftarrow N^c$ (where N^A is a set of candidate nodes)
9. If token > N^c
   Set higher voting weight
   Else
   Voting weight is low
10. Nodes with higher votes validate transactions
11. Get the reward
12. Stop

Tendermint

Tendermint is designed to resolve the issues of PoW which are speed, scalability, and environmental issues. The tendermint handles 33% of the byzantine actor in the blockchain network. It is supported by any programming language. Validators perform a major role in the network, and they took part in validating every transaction and committing new blocks in the blockchain. Validators are taking part in a consensus algorithm to broadcasting a cryptographic signature which is used to act as a vote to extend the blockchain. Whoever wants to become a validator in the tendermint, then user holds tendermint network for sometimes and locked it as voting power. Tendermint is working based on the delegation concept. The user or delegators of the system put their staking token at stake with a validator of choice.

Table 2 Comparison of different consensus algorithms

Property	PoW	PoS	PBFT	DPoS	Tendermint
Speed	Slow	Normal	Fast	Fast	Normal
Energy consumption	High	Medium	Low	Partial	Low
Security	Secure	Secure	Least Secure	Secure	Secure
Degree of centralization	Very low	Low	High	Low	High
Blockchain type	Open	Open	Permissioned	Open	Permissioned
Transaction rate	Low	High	High	Medium	High
Transaction finality	Probabilistic	Probabilistic	Immediate	Not fixed	Not fixed
Scalability	High	High	Low	Medium	High
Trust model	Untrusted	Untrusted	Semi-trusted	Untrusted	Semi-trusted
Byzantine fault tolerance (%)	50	50	33	50	50
Performance (latency)	High	High	Excellent	Medium	High
Node management	Open	Open	Permissioned	Open	Permissioned
Token needed	Yes	Yes	No	Yes	NA
Energy saving	No	Partial	Yes	Yes	Yes
Performance (throughput)	Limited	Limited	Excellent	Limited	Limited
Tolerated power of advisory	$<=25\%$ computing power	$<51\%$ stake	$<=33.3\%$ fault replicas	$<=51\%$	$<33.3\%$ Byzantine voting power
Example	Bitcoin	Peercoin, Ethereum	Hyperledger fabric	Bitshares	Tendermint

The delegator has some losing tokens if the validators not working according to specified protocol rules. At least four validators are working in the tendermint network but there are no maximum limits of validators. Cosmos project is using tendermint consensus algorithm which has 100 validators but later it will be increased by 300 validators. For finalizing block, it required 3 s but it can be achieved within a second as well. The finalization of the block depends on the number of validators (Table 2).

Tendermint follows the following algorithm

1. Start
2. Proposed new block
3. If a proposed block is valid
   Go to step 4
   Else
   Go to the pre-vote stage
4. Proposed block goes to the pre-commit mode
5. Wait for pre-votes from + 2/3
6. If pre-commit is found
   Go to commit state
   Else
   Create a new height of block (wait until "commit-time" + "delta")
7. Get block in commit state
8. Save and stage block
9. Broadcast the commit
10. Get + 2/3 commits
11. Set "commit-time"
12. Stop

4 Conclusion

Blockchain technology is trending in the market, and different sectors are adopting blockchain technology. In the blockchain, block creation is achieved by multiple steps.

This paper explains a few consensus algorithms in detail and examines the requirements of consensus algorithms in blockchain technology. The many blockchain platforms are used to categorize the various blockchain consensus algorithms. Public, private, and consortium blockchain platforms are the three types of blockchain platforms. Different consensus algorithms are built based on the parameters of each blockchain type. The study also compares and describes some prominent consensus algorithms, including PoW, PoS, PBFT, DPoS, and tendermint. The study also explains how the consensus methods function step by step. The comparison of consensus algorithms shows that PoW and PoS maintain higher scalability as compared to other consensus algorithms discussed. Performance-wise DPoS and PBFT are faster. PoW consumes high energy, whereas PBFT, DPoS, and tendermint consume the least energy. DPoS and PBFT are much faster than PoW.

References

1. Bach LM, Mihaljevic B, Zagar M (2018) Comparative analysis of blockchain consensus algorithms. In: 2018 41st International convention on 8 information and communication technology (MIPRO), ISBN 978-1-5386-3777-7, pp 1545–1550, 21 May 2018. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) Conference 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).
2. Samsir N, Jain A (2021) Overview of blockchain technology and its application in healthcare sector. In: 6th International conference on information and communication technology for competitive study, Dec 2021
3. Gemeliarana IG, Sari RF (2018) Evaluation of proof of work (POW) blockchains security network on selfish mining. In: 2018 International seminar on research of information technology and intelligent systems (ISRITI), ISBN 978-1-5386-7423-9, pp 126–130, Nov 2018
4. Nguyen GT, Kim K (2018) A survey about consensus algorithms used in blockchain. J Inform Process Syst 14:101–128
5. Yang F, Zhou W, Wu Q, Long R, Xiong NN, Zhou M (2019) Delegated proof of stake with downgrade: a secure and efficient blockchain consensus algorithm with downgrade mechanism. Special Section on Emerging Approaches to Cyber Security IEEE, ISSN 2169-3536, vol 7, pp 118541–118555
6. Jalal I, Shukur Z, Azmi KA (2020) A study on public blockchain consensus algorithms: a systematic literature review. Mathematics and Computer Science, ISSN 10.20944, vol 1 issue 1, Nov 2020
7. Hasselgren A, Kralevska K, Gligoroski D, Pedersen SA, Faxvaag A (2020) Blockchain in healthcare and health sciences-A scoping review. Int J Med Inform, ISSN 1386-5056 134
8. Chaudhry N, Yousaf MM (2018) Consensus algorithms in blockchain: comparative analysis, challenges, and opportunities. In: 2nd International conference on open-source systems and technologies (ICOSST), 978-1-5386-9565-4, 19th Dec 2018, pp 54–63
9. Mingxiao D, Xiaofeng M, Zhe Z, Xiangwei W, Qijun C (2017) A review on consensus algorithm of blockchain. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC), intelligent sustainable systems, ISBN 978-1-5386-1646-8 vol 2, pp 2567–2572, Oct 2017; Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
10. Wang S, Ouyang L, Yuan Y, Ni X, Han X, Wang FY (2019) Blockchain-enabled smart contracts: architecture, applications, and future trends. IEEE Trans Syst, Man, Cybernetics: Syst, ISSN-2168-2232 49:2266–2277

Attack Detection in Internet of Things: A Systematic Literature Review

Jyoti Mante and Kishor Kolhe

Abstract As the frequency of security breaches continues to rise, cybersecurity remains a critical concern for every industry in the online. Thousands of zeroday attacks are known to emerge on a regular basis as a consequence of the integration of multiple protocols, primarily from the Internet of Things (IoT). The majority of such attacks are minor variants on previous research findings intrusions. This suggests that even sophisticated techniques like typical machine learning (ML) algorithms have difficulties spotting these tiny kinds of attacks over time. These attacks are called as DDoS attack; they are used to prevent clients from accessing a server or website. DDoS attacks have been employed by cybercriminals to bring down targeted servers and breach venture networks with the ability to overwhelm results. Because of the growing volume and complexity of DDoS attacks, many organizations are having difficulty handling them. Smart gadgets and IoT are particularly vulnerable to a wide range of DDoS hits due to resource constraints such as limited memory and processing capacity, thus cybercriminals are aware of these current technologies and their flaws. Because of an attack on their internet service providers in 2016, many firms, including Netflix, CNN, and Twitter, were forced to go down for nine hours. This technological failure resulted in several issues, including financial losses, productivity losses, brand damage, insurance rating drops, unstable client-provider relationships, and IT budget overruns. We need to construct an IDS system to expose and prevent DDoS attacks to secure data processing, information technology, and commercial components. The cost of cybersecurity will be greatly lowered if security teams use current and new technology like ML, automation, and artificial intelligence (AI). This study will examine the detection performance of DDoS attacks using several ML, DL techniques and also categorize it into cloud and fog ecosystems.

Keywords Cyber security · DDOS attack · Machine learning · IoT · Fog networks

J. Mante (✉) · K. Kolhe
School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India
e-mail: jyoti.khurpade@mitwpu.edu.in

K. Kolhe
e-mail: Kishor.kolhe@mitwpu.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_24

1 Introduction

When attackers send huge numbers of requests to overload the victim's system resources, they are committing a DDoS attack [1, 2]. In order to convey malicious payloads, attackers often target vulnerabilities in application, transport and network layer protocols [3, 4]. With the fast expansion of Internet access to smart devices and the IoT, which is expected to reach 500 billion people by 2030 [5], poorer countries are becoming more concerned. We provide a comparison of multiple ML algorithms, both classic ML-based and more sophisticated DNN-based, that are utilized to protect our networks against DDoS attacks.

1.1 IOT and Security

The IoT has grown in popularity tremendously over the years, and Gartner has identified it as essential technological trends for 2021. According to analysts, by the end of 2022, more than 43 billion linked IoT devices will be installed and more than 75 billion devices by the end of 2025 [6, 7]. As per CISCO, Machine-to-Machine (M2M) connectivity will increase from 6.2 billion in 2018 to 14.9 billion in 2023, representing a 19% rise yearly.

On the other hand, the IoT revolution creates new opportunities, but it has the potential to cause substantial disruption, particularly in terms of security and privacy. One of the most frequent attack vectors nowadays is IoT botnets, which are made up of compromised IoT devices and are established by taking advantage of the ubiquitous accessibility of IOT connected devices, as well as their lack of security. The main motives for running IoT botnets is to launch large-scale DDoS attack via IoT endpoints, among other things.

Mirai attack in 2016 infected 600,000 unsecured IoT devices and exploited them to attack Dyn, a DNS provider, with a peak traffic rate [6] of 1.2 Terabits-per-second (Tbps).

1.2 Cloud Computing and Fog Computing

Cloud computing (CC) is strongly linked to the IoT. The CC paradigm relies on centralized computing and is hampered by latency and bandwidth issues. As a result, interest in novel decentralized computing models, such as Fog computing, is growing.

FC is a distributed architecture that emerged from the necessity to handle the challenges that the IoT revolution has brought to CC. CISCO [1] was the first to propose FC. It serves as a form of middleware that connects the Cloud with the IoT, supporting scalability, interoperability, latency, location awareness, and security [4] (Fig. 1).

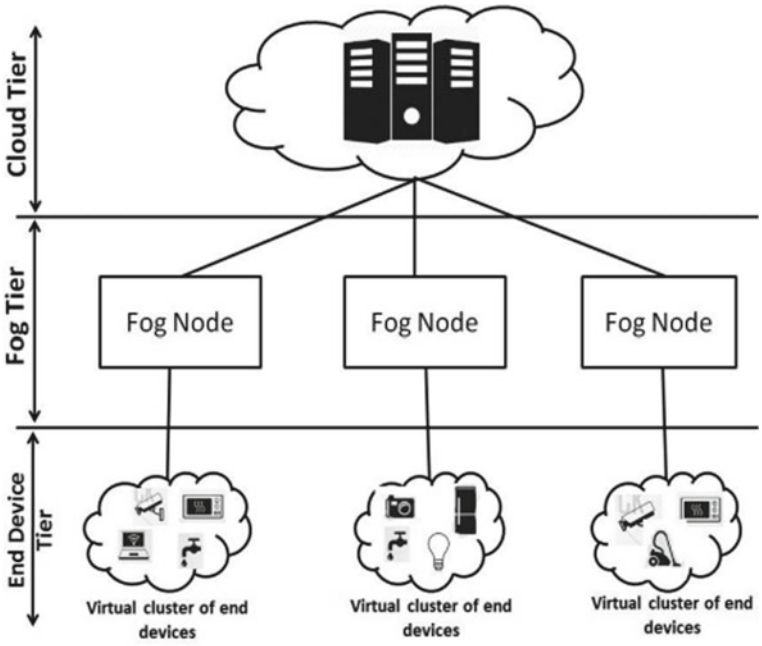


Fig. 1 Fog computing architecture

The security problem of FC cannot be overlooked because of large distribution, limited computing and storage resources [8]. Because of these factors, the fog node is particularly vulnerable to outside attacks [9], such as DDoS, R2L, U2R, PROBE and so on.

Although DL techniques have shown to be better in distributed attack detection, executing the attack detection model at the fog layer in a real-time application is problematic owing to the restricting constraints of fog nodes. As a result, detecting, mitigating, and preventing DDoS attacks in a timely and efficient manner remains a difficult challenge [10]. Because of the substantial annual increases in cyberattacks, ML technologies are being used to help combat the growing threat of cyberattacks [6].

Section 2 of this work contains the extensive literature review, and Sect. 3 gives the results and analysis phase that has been evaluated in the literature study. Result and discussion section gives the detail overview of several research employing Machine Learning and Deep Learning techniques.

2 Literature Review

2.1 Literature Survey

Rathore and Park [1] developed a unique fog detector utilizing ELM Semi-Supervised Fuzzy C-Means (ESFCM) NIDS. For limited resource and network detection, this distributed IDS uses fog computing to handle geographically distributed and low-latency IoT detection. The authors proposed an ESFCM classification that combines the Supervised Fuzzy C-Means (SFCM) and ELM classifiers to identify known and unexpected threats faster.

Diro and Chilamkurti [2] proposed an IDS for IoT systems based on FC. A pre-trained stacked auto encoder is utilized for feature engineering, while softmax function is employed for classification. Fog nodes are used to detect attacks. But the system has several drawbacks as it uses the approach of SGD combined with the backpropagation method for classification which takes more time for updating and training the models also system complexity is more.

Sundarakantham et al. [3] developed an online sequential extreme learning machines-based FC detection method (OS-ELM). The attacks are detected at local fog nodes and the state of the IoT system updated at a cloud server. With failure of centralized server complete system will stop working. Also the detection model suffers from low detection, accuracy, and high false alarm rate. To overcome this drawback Kumar [5] proposed leveraging FC to safeguard IoT networks using a distributed ensemble design-based IDS. To increase the overall detection rate, an anomaly detection is performed which uses a RF based ensemble method. Then the regular traffic is allowed to be utilized by the IoT system, but dangerous behaviour are notified to the admin. The log information for such anomalous traffic is stored at cloud server, and an effective intrusion response mechanism can be initiated based on it.

Amma and Selvakumar [4] proposed a vector convolutional DL-based anomaly detection framework. Their goal was to overcome the non-scalability of IoT traffic anomaly detection systems. The IoT traffic properties were dispersed over many fog nodes for a robust anomaly detection system. Also, The FCN was used to train the VCDL network. Priyadarshini and Barik [11] developed a DL method for DDoS protection for FC using SDN. The SDN network is employed, with a DL-based DDoS protection module and LSTM model was selected because it performs better with sequential data as LSTM algorithm improves accuracy and reduces failure rates. The system proposed by Ahmed et al. [12] uses the LSTM DL framework and is modified via a CC layer service. During the training using LSTM model, the edge layer data is collected and transfer to the cloud. The trained model served as a classification engine on the fog layer nodes. The system's flaw is that it must detect edge data to train the LSTM network on the cloud.

To enable faster and more accurate anomaly detection, Lawal et al. [8] presented a hybrid anomaly mitigation system for IoT that uses FC. To enable speedier attack detection, the signature-based module uses an IP blacklist, the anomaly-based

module, on the other hand, employs an XGB approach to identify network traffic flow as normal or abnormal. Using fog-based attack detection, Alrashdi et al. [13] have introduced the FBAD framework, an effective ensemble approach for detecting abnormal activity in IoT health-monitoring system. To detect attacks on fog-based IDS, this system employs an ensemble of OS-ELM algorithms to identify normal and attack records. The framework is suggested in a fog layer to overcome the limitations of cloud-based detection.

Diro and Chilamkurti [14] created a distributed attack detection solution based on DL for IoT framework. The authors proposed this technology to host attack detection systems and training models at the FC layer. Because they are near to the smart freamwork provided by the social IoT, Fog nodes are edge nodes that host threat detection systems and train models. Almiani et al. [15] suggested an intelligent IDS based on multi layered RNN having two main engines: traffic analysis and classification.

Pahl and Aubet [10] developed an IoT anomaly detection system employing K-means clustering and BIRCH. The suggested approach is used to develop site-invariant IoT S models and a periodicity mining technique.

FlowGuard is a protection mechanism proposed by Jia et al. [16] against IoT DDoS threats for detecting, identifying, classifying, and mitigating IoT DDoS threats. Authors introduce DDoS attack detection technique based on traffic fluctuations, as well as two ML models for DDoS detection and categorization. Pajouh et al. [17] used a recurrent neural network to create an intrusion detection technique for detecting IoT malware. Three stages are used in the proposed mechanism: data gathering, feature extraction, and deep threat classification. Hong et al. [6] proposed research on DDoS detection in the IoT setting. They suggested a 7-layer sequential model for DDoS attack detection and proposed a novel classifier for feature extraction. They've also included number of criteria to prevent ML-based DDoS detection from misclassifying regular traffic as DDoS traffic. Author Ibrahim et al. [18] proposed a multilayered system for detecting Botnets used in DDoS attacks that uses ML methods. They used feature extraction, classification, and hyper-parameter setup for improving the performance of system (Table 1).

3 Result and Discussion

The specifics of the experiment, including the dataset utilized, performance parameters, analysis of results, and discussion are presented in this section.

3.1 Dataset Description

The researchers used a range of datasets such as CICIDS-2017, CICIDS-2019, NSL-KDD [19] and BoT-IoT to evaluate their methods.

Table 1 Summary of the various techniques used in DDoS attack detection

Paper	Techniques	Dataset	Performance (Acc in %)	Strategy	Framework
[1]	Fuzzy C-Means, ESFCM	NSL-KDD	86.53	Centralized	Cloud computing
[2]	DL Techniques, Shallow Learning, Deep Model	NSL-KDD	99.20	Distributed	Fog computing
[3]	NB, ELM, ANN, OS-ELM	NSL-KDD	97.36	Distributed	Fog computing, Cloud computing
[4]	VCDL approach, CNN	Bot-IoT	99.98	Distributed	Cloud computing
[5]	K-NN, XGBoost, and NB	NSL-KDD, DS2OS	99.99	Distributed	Fog computing
[11]	LSTM	Hogzilla	98.88	Centralized	Fog computing
[12]	DNN, LSTM, Bi-LSTM, GRU, CNN, CNN-LSTM	KDDCUP99, UNSW-NB15, CICIDS-2017, NSL-KDD	99.97	Distributed	Fog computing
[8]	XGBoost, DT, k-NN, NB, GRB	BoT-IoT	99.00	Distributed	Fog computing
[13]	DT, NB, SVM	NSL-KDD	98.51	Centralized	Fog computing
[14]	SGD	NSL-KDD	99.00	Distributed	Fog computing
[15]	DRNN	NSL-KDD	98.21	–	Cloud computing, Fog computing
[16]	LSTM, CNN	CIC-DDoS2019	99.90	–	Cloud computing

3.2 Result Analysis

To demonstrate the success of the system, accuracy parameter is employed in the assessment process of the various proposed models used by researchers in centralized and in distributed IoT system at cloud and fog level. Table 2 compares the accuracy of the ML, DL models and Fig. 2 shows the accuracy comparison of those techniques, from comparison of models we can say that researchers at [5, 12] shows better accuracy of 99.50% + while [1, 20] performs poor compare to other related techniques as their accuracy is below 85.0%.

Table 2 Accuracy comparison of various researcher's techniques used to detect DDOS attack

Paper (Best performing algorithm)	Accuracy	Paper (Best performing algorithm)	Accuracy
[1] ESFCM	86.53	[14] SGD	99.00
[2] DEEP MODEL	99.20	[20] CF-KNN	84.66
[3] OS-ELM	97.36	[15] DT, RF, ANN	99.4
[4] VCDL	99.98	[10] DRNN	98.21
[5] XGBOOST	99.99	[16] KMEANS	96.3
[11] LSTM	98.88	[17] CNN	99.90
[21] CORR	86.53	[22] LSTM	98.80
[12] CCN-LSTM	99.97	[6] BLSTM-RNN	99.00
[8] XGBOOST	99.00	[18] LPA	99.99
[13] DT	98.51	[7] KNN	95.91

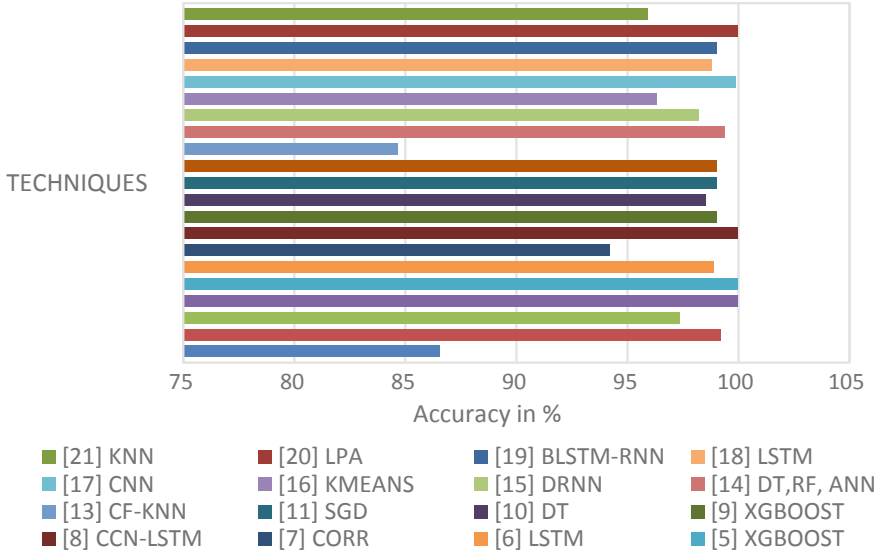


Fig. 2 Accuracy comparison of various researcher's techniques

3.3 Limitations and Future Scope

Signature detection shows limit on detecting new forms of threats. Anamoly detection generates high false positives and doesn't give clear information about the malicious events in some forms of attacks. Feature extraction and selection techniques used decreases the accuracy.

So a work need to be done in future to detect the DDoS attacks more accurately at fog nodes as more IoT devices data will be generated in future. An efficient decision

making machine learning model need to be developed using reinforcement learning for efficient DDos attack detection and prevention.

4 Conclusion

DDoS attack pose a severe threat to many aspects of our lives. To reduce the number of attacks, we need a sophisticated intrusion detection system. This analysis utilizes DARPA, CICD2017, CICDoS2019, NSL-KDD, BoT-IoT. It also looked at ML and DL algorithms. Performance evaluation uses accuracy metrics. The research also stated that the XGBOOST and Hybrid DL algorithm (CNN-LSTM) had the best accuracy around 99.5%. Our work attempts to contribute to this field's research. Hybrid deep learning, deep learning and Machine Learning algorithms gives best accuracy of compare the approaches in detecting attack traffic.

References

1. Rathore S, Park JH (2018) Semi-supervised learning based distributed attack detection framework for IoT. Appl Soft Comput 72:79–89
2. Diro AA, Chilamkurti N (2018) Deep learning: the frontier for distributed attack detection in fog-to-things computing. IEEE Commun Mag 56(2):170–175
3. Sundarakantham K, Prabavathy S, Shalinie SM (2018) Design of cognitive fog computing for intrusion detection in the internet of things. J Commun Networks 20(3):291–298
4. Amma BNG, Selvakumar S (2020) Anomaly detection framework for Internet of things traffic using vector convolutional deep learning approach in fog environment. Future Gener Comput Syst 113:255–265
5. Kumar P, Gupta GP, Tripathi R (2021) A distributed ensemble design based intrusion detection system using fog computing to protect the internet of things networks. J Ambient Intell Humanized Comput
6. Wehbi K, Hong L, Al-salah T, Bhutta AA (2019) A survey on machine learning-based detection on DDoS attacks for IoT systems. In: 2019 Southeast Conference. IEEE, pp 1–6
7. Koroniotis N, Moustafa N, Sitnikova E, Turnbull B (2019) Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset. Future Gener Comput Syst 100:779–796
8. Lawal MA, Shaikh RA, Hassan SR (2020) An anomaly mitigation framework for IoT Using fog computing. Electronics 9:1565
9. Pajouh HH, Javidan R, Khayami R, Ali D, Choo KKR (2019) A two-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks. IEEE Trans Emerg Topics Comput
10. Pahl M-O, Aubet F-X (2018) All eyes on you: distributed multi-dimensional IoT microservice anomaly detection. In: Computer science 2018 14th international conference on network and service management (CNSM), pp 72–80
11. Priyadarshini R, Barik RK (2019) A deep learning-based intelligent framework to mitigate DDoS attack in fog environment. J King Saud Univ—Comput Inf Sci
12. Samy A, Yu H, Zhang H (2020) Fog-based attack detection framework for internet of things using deep learning. IEEE Access 8

13. Alrashdi I, Alqazzaz A, Alharthi R, Aloufi E, Zohdy MA, Ming H (2019) FBAD: fog-based attack detection for IoT healthcare in smart cities. In: 2019, IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON)
14. Diro AA, Chilamkurti N (2017) Distributed attack detection scheme using deep learning approach for Internet of Things. Future Gener Comput Syst
15. Almiani M, AbuGhazleh A, Al-Rahayfeh A, Atiewi S, Razaque, A (2020) Deep recurrent neural network for IoT intrusion detection system. Simul Model Pract Theory 101:102031
16. Jia Y, Zhong F, Alrawais A, Gong B, Cheng X (2020) FlowGuard: an intelligent edge defense mechanism against IoT DDoS attacks. IEEE Internet Things J 7(10)
17. HaddadPajouh H, Dehghantanha A, Khayami R, Choo KKR (2018) A deep recurrent neural network-based approach for the internet of things malware threat hunting. Future Gener Comput Syst 85:88–96
18. Ibrahim WNH et al (2021) Multilayer framework for botnet detection using machine learning algorithms. IEEE Access 9:48753–48768
19. NSL-KDD Dataset. https://www.unb.ca/cic/datasets/nsl.html. Accessed 15 Feb 2019
20. Hasan M, Islam MM, Zarif MII, Hashem M (2019) Attack and anomaly detection in IoT sensors in IoT sites using machine learning approaches. Internet Things 7(2019):100059
21. Ghazy RA, El-Rabaie E-SM, Dessouky MI, El-Fishawy NA, Abd El-Samie FE (2020) Feature selection ranking and subset based techniques with different classifiers for intrusion detection. Wirel Pers Commun 111(1):375–393
22. McDermott CD, Majdani F, Petrovski AV (2018) Botnet detection in the internet of things using deep learning approaches. In: 2018 International joint conference on neural networks (IJCNN)

Effect of COVID-19 on Construction Industry in Mumbai

Hemant Kothari, Divya Hiran, Himanshi Panwar, and Shivoham Singh

Abstract The pandemic has hit many businesses from March 2020.This effect remains over a year and may continue further depending on the efficacy of the vaccine and the variants produced by the virus. The construction industry which is completely dependent on labour (unskilled, semi-skilled and skilled workers) has had an impact. We studied this impact and analysed the main causes of the damage. This work is on two incidences, the first and the second lockdown. Areas selected in Mumbai city. The sector has been divided in three verticals, Real Estate (Organized), Real Estate (Un-Organized) and the Supplier Segment. Our study reveals an impact due to three main reasons, labour movement, material cost and lock-down restrictions. It is clearly visible that the movement of labour is better in the organized sector than the un-organized and supplier segment. This leads to the fact that we as a country, are moving towards industrialization of the construction segment, where work force is preferring facilities and benefits over pay. A positive change displayed in our study and the process towards westernization is certainly a positive way forward.

Keywords Construction industry · Effect of COVID-19 · Infrastructure · Real estate

1 Introduction

The Indian economy is developing at fast pace lead by the construction industry. Construction industry play a pivotal role in overall growth of all the industries in the economy. The basic infrastructure for the industries whether they are in production of goods or delivering services are backed by the construction Industry. The Indian construction sector is the fuel for the country's overall development.

H. Kothari · H. Panwar · S. Singh (✉)
Pacific Academy of Higher Education and Research University, Udaipur, India
e-mail: shivohamsingh@gmail.com

D. Hiran
Department of Home Science, Govt. Meera Girl's College, Udaipur, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_25

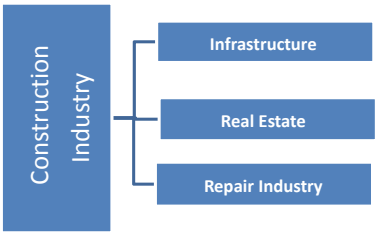


Fig. 1 Construction industry

"According to Invest India, the expected growth rate of the Indian construction industry would be 15.7% and reach $738 billion by 2022. Construction industry contributes 55%, 15% and 30% shares, respectively, in steel, paint and glass industries.

According to the data of Department for Promotion of Industry and Internal Trade (DPIIT) the foreign Direct Investment (FDI) in construction industry was $ 25.66 billion during April 2020 and September 2020. For infrastructure development there is need of $777 billion in India for the development of infrastructure across sectors by 2022 as per the government of India estimates.

By 2022, India is expected to become the world's 3rd largest construction market."

The construction industry in India consists of three verticals, infrastructure, real estate and repair (Fig. 1).

1.1 Infrastructure

A magazine on construction industry (Civil Engineering and Construction Review) stated that The National Highways Authority of India (NHAI) is planning to award projects worth around Rs. 2.25 lakh crore in the current fiscal. The authority awarded 141 projects (4788 km) worth Rs. 1.71 lakh crore in 2020–21.

1.2 Real Estate

The magazine (Civil Engineering and Construction Review) also gives this information on the real estate sector. The housing sector growth in next few quarters is going to support the economic growth of the country. The demand for the housing sector will get boost from both those who are buying first-time and those who are moving up in the property ladder by shifting in bigger homes or acquiring another one.

According to India Brand Equity Foundation (IBEF) recent report, the real estate sector in India with present 6–7% contribution to the country's GDP, is expected to reach a market size of $ 1 trillion by 2030 and contribute 13% to the country's GDP by 2025.

1.3 Repair Industry

This sector includes the repair of old buildings, highways, revamp of airports and other structures. Traditionally, the repair sector has always shown a steady double digit grow year on year. The demand for repair will always exist as the projects age with time. Since this is quiet an unorganized structure, accurate data on the market size is not available. With my experience in the industry, we estimate that this market is approximately 30–40% of the new market by size.

In the last decade and coming few years, India has been and will be very active in this segment. The report from Niti Ayog explains the need of roads, railways, sea ports and airports in our country. This infra will generate internal growth and also invite foreign investments. Real estate development is also extremely important to the growing needs of our population. Influx of rural movement to the urban territories has always been on the higher side.

Our study limits to the Mumbai city and surroundings. This is a very large market for construction. The sample has been split into three groups as below,

a. **Organized Sector**

   In this sector, we have taken reputed companies having multiple projects in the city. They are large organizations who have all employee systems in place. They offer standard facilities to the labour including basic medical help, decent labour camps and employee state insurance facilities. Their sub-contractors (labour contractors) have all the licenses required. These organisations have a HR or IR department who insure that the right practises are followed.

b. **Un-Organized Sector**

   This is a larger group by numbers and they have mostly single projects going on in the city. They are smaller organisations who engage labour contractors for their projects. They do not have any formal HR or IR department and the compliance to formal licenses or facilities or medicals may be complete. The benefits offered to the labour are limited as compared to the organized sector.

c. **Supplier Sector**

   This is a different group of organisations who do not build anything but play a major role in both the sectors. They are the suppliers of building materials. We have clubbed the manufacturers and traders in this sector because both have an equal role to play. This sector is a support function and their opinion certainly matters for our study.

2 Literature Review

Majumder and Biswas in his study highlighted the impact of COVID-19 in the construction industry along with the risk assessment. They also discussed the implementation of the safety measures during the pandemic for the workers.

Zamani et al. [1] in their study discussed the financial and operational problems triggering due to the COVID-19 accompanied by the suggestions to overcome from the crises.

Studied a publication by researcher Pathak [2]. She writes about the workers in the SME sector which are very similar to the construction sector. Her techniques help in improving the efficiency of the workers in the SME sector. Construction workers also are part of the SME because they work for a contractor which is always a small or medium or sometimes even micro unit. Even in very large infrastructure projects, the work is distributed by the main contractor to such independent small contractors. However, there is a difference between the workers who are in a manufacturing set up when compared to the same in construction sector. Manufacturing is long term whereas construction (for both infra and real estate) are short term projects. The construction workers will always stay on the site and labour colony will have to be made for them. The manufacturing sector ensures that their workers will have a permanent resident. The manufacturing sector workers once a year (mainly monsoon) visit their home which is mostly a rural location and sometime indulge in farming for extra income.

Any other paper specifically mentioning effect of COVID on the construction segment in India was not traced by us.

3 Research Methodology and Observations

The primary objective of the study was to know the Effect of COVID-19 on Construction Industry in Mumbai.

To meet the objective of the study the population was divided in three sub sections of the industry namely organised sector, un-organised sector and supplier sector. The purposive quota sampling was used to collect the data from each group. The survey method was used to collect the data. Twenty telephonic interviews were conducted in each of the subgroups (Fig. 2).

The questions were asked related to the size of the impact, cause of the impact and movement of labour in the two lock-downs which happened in the year starting from March 2020 to June 2021 (Table 1).

Fig. 2 Sample design

Table 1 Perception regarding effect of COVID-19

	Strongly agree	Agree	No effect	Disagree	Strongly disagree
Organized sector	16	4	0	0	0
Un-organized sector	19	1	0	0	0
Supplier (Building material)	18	2	0	0	0

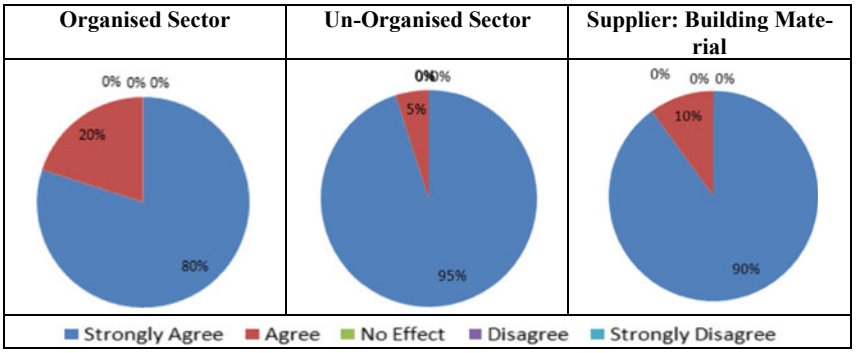


Fig. 3 Perception regarding effect of COVID-19

1. The size of the impact was strong and we had the majority of sample saying it. The graph of all three sectors denote the same message. Overall, the effect is felt stronger in the un-organized and supplier segment (Fig. 3).

2. The reason for impact. The survey reveals three prime reasons for the covid impact. Labour movement, high rise in the material cost (cement and steel) and lockdown which restricted the movement of manpower and materials. Work from home, does not function in the construction sector and thus lockdown effect was more significant in this sector (Fig. 4).

3. The labour movement in this impact has been observed to follow a trend. We have observed that the recovery of the organized sector was far better than that of the un-organized and supplier sectors. This shows that the labour that was not in the organized sector did not completely return or has moved towards the organized sector (Fig. 5).

In the first lockdown, the organized sector was a 5% negative recovery but in the second lockdown, they got a 16% higher return of the work force. These calculations show that they got more labour than they had before in the second lockdown (Fig. 6).

In the un-organized sector, the recovery on first lockdown was positive, but the second lockdown had a deficit in the labour recovery. Similar was the result in the supplier sector (Fig. 7).

We also asked all the three sectors on what they felt about the future of the construction sector. The results are compiled as below (Fig. 8),

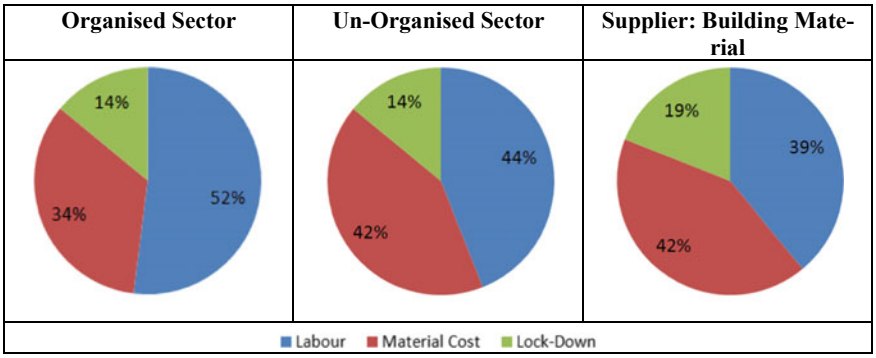


Fig. 4 Reason of impact effect of COVID-19

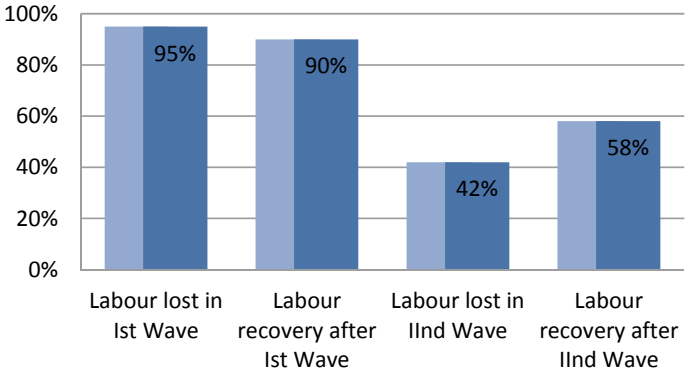


Fig. 5 Labour movement: organized sector

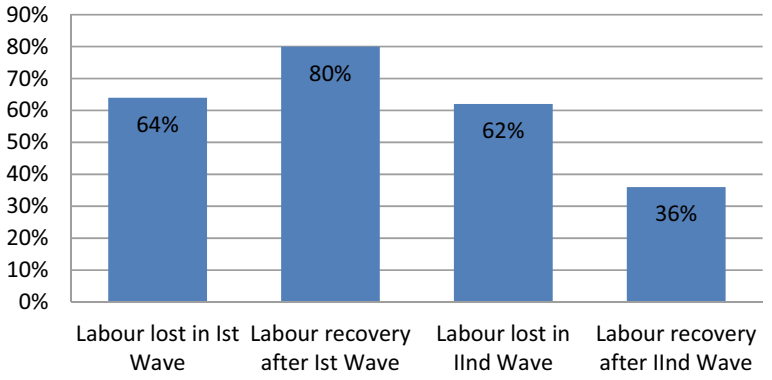


Fig. 6 Labour movement: un-organised sector

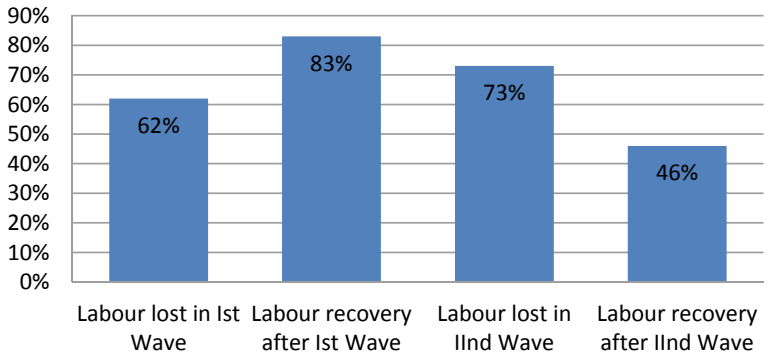


Fig. 7 Labour movement: building material supplier sector

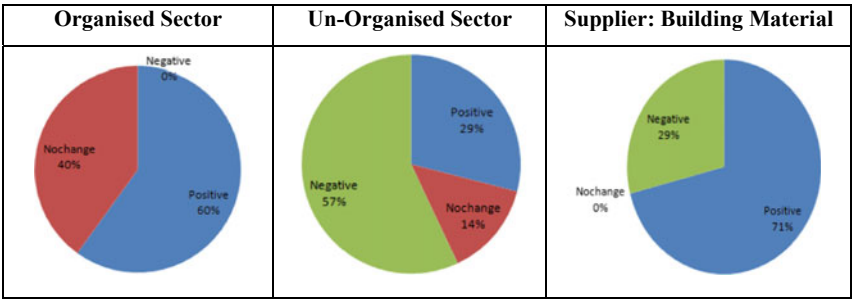


Fig. 8 Future outlook of construction industry

Fear looms the un-organized sector as we can see more negativity. The supplier sector is supplying to both and have a larger group positive about the future. The organized sector has 40% no change and 60% positive. This is connected with the recovery of labour, and it is very evident that they will do better than the un-organized sector.

3.1 Effect of Covid-19 on Construction Industry

- High cost of commodities (Steel, aluminium, and copper)
- Liquidity crisis due to high levels of debt.
- Renegotiation of project contracts.
- Safety of own people
- Supply chain disruption
- High labour costs
- Low market liquidity
- Failure of subcontractors
- Increased raw material costs
- Increased cost due to COVID protocols.
- High cost of specialised equipment and materials.

4 Conclusion and Future Scope

Based on our study, we can clearly conclude that the movement of labour is more towards the organised sector. They can witness better working conditions and facilities which are provided including medical benefits which have shown great importance in the pandemic. As in developed countries, the culture of benefits in construction and industrial sector are same. We call this "Industrialization of Civil Engineering." India is also growing towards this industrialization in the construction sector and evident on the movement of labour.

We can also see the escalation of material cost (steel almost double and cement up by almost 50%) is a major reason for the impact to construction. The rise in construction cost cannot be passed on to the home buyers because they are already in the financial pressure created by the pandemic. The Maharashtra government gave benefit in the stamp duty, but it was a small effect.

The lockdown restricted officer level staff entering the project site and was also a deterrent to the construction sector. Restriction of movement helped the pandemic but certainly did not favour the construction sector.

Scope for Future

This study can be done in other major cities such as Delhi, Kolkata, Chennai, Bangalore, and Hyderabad to see if similar effects have been observed.

We are sure that the organized sector is a better recruiter and a deep study can be done on what they have to offer which has suddenly developed more significance post pandemic.

References

1. Zamani S, Rahman RA, Fauzi M, Yusof L (2021) Effect of COVID-19 on building construction projects: impact and response mechanisms. IOP Conf Ser: Earth Environ Sci 682:012049. https://doi.org/10.1088/1755-1315/682/1/012049
2. Pathak K (2020) Evaluating efficiency of human resource management strategies for developing micro small and medium enterprises. Dayalbagh Educational Institute. http://hdl.handle.net/10603/327471

TranslateIT: Android-Based Mobile Application for Multilingual Translation

Ishan Somani, Deevesh Chaudhary, Devesh Kumar Srivastava, and Deepika Shekhawat

Abstract Over the years, the issue of linguistic variation has posed a significant barrier to successful information collaboration. In modern times, learning multiple languages can be both time consuming and a hectic process because of the language difference, and it can also be an expensive expenditure if a tutor is appointed for teaching the same. Language interpreters must be able to communicate in and understand both languages, because the traditional approach to handling the problem of translation is neither productive nor advantageous. TranslateIT is a mobile application that can be used to translate among 11 native languages that includes English, Afrikaans, Arabic, Belarusian, Bulgarian, Bengali, Catalan, Czech, Hindi, Urdu, and Welsh. Therefore, we have developed an android-based mobile application called TranslateIT which is a multi-language translator which has been made using firebase ML in Android Studio using Java. Because of this application, tourists will be able to communicate more effectively with locals and gain access to more relevant information, which will help them learn more about their surroundings.

Keywords Android · Firebase ML · Language translator

1 Introduction

Because of the ever-increasing use of mobile devices, the concept of mobile computing and universal access has become a very important part of our day-to-day lives. This is directly because of the growing processing capacity, large storage space, simple user understanding, and improved network structure that has been made possible by the increasing use of mobile gadgets. There is a growing number of mobile applications that support our daily activities and provide a variety of leisure options. Android is unquestionably the most popular app available today, being used by millions of smartphones and tablets and increasing at an exponential rate. As a result, the Android phone is the most technologically innovative and

I. Somani · D. Chaudhary (✉) · D. K. Srivastava · D. Shekhawat
School of Information Technology, Manipal University Jaipur, Jaipur, Rajasthan, India
e-mail: choudhary.deevesh@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_26

user-friendly products available. Translation and language learning technologies are required to benefit from a modern education system. On many occasions, the current Internet's functionality is reached its maximum, and there has been a large increase in the number of mobile applications that supply those services. Translation into another language by machine is referred to as machine translation (MT). Arithmetical machine translation (AMT) is a form of machine translation that has been classified using machine learning methods.

Human beings have attempted to find a way out to the problems of language interpretation for hundreds of years now, and they have been successful in some cases. Over the course of a century, humans have devised a variety of methods for translating languages to address the issues connected with linguistic disparities. A human translator who is capable of understanding and translating both languages was employed as the initial step in resolving the language issue. This strategy proved to be successful and was adopted by the other parties involved. As the first way ever introduced and utilized for decades, this method has shown to be less efficient and successful than other methods of language translation that have been discovered through research across the years. According to this method, the translator must be able to understand and articulate the language into which the translation is being performed, as well as understand and speak about the language to which the translation is being performed. Foreign language translation is beneficial in many areas, including educational settings. When the individuals do not comprehend the language of the tutor, it is difficult to teach in a specific language. An interpreter will be required in order for the pupils to gain a comprehensive comprehension of what they are being instructed in. In the field of tourism, travelers may find themselves unable to speak effectively with the people in the tourist country that they have visited, hence limiting contact.

As a result of misconceptions in common language, the spread of erroneous information may be prohibited altogether. Due to the immense cultural diversity of some countries, such as India, where 121 languages are spoken, language comprehension is a critical aspect in political decision-making. As a result, each representative should be able to successfully interact with the other in order to discuss their racial viewpoints with one another. Language comprehension is another key component in the entertainment industry since, in order for viewers to grasp any content related to entertainment, they must be familiar with the breadth of language used in a variety of different fields. Language is an essential component of interaction; without it, it is impossible to achieve meaningful results in this situation. Consequently, language translation is extremely crucial for the organization, regardless of the field in which it operates. For this reason, it is critical to create an alternative to the standard translation of human languages by telephone, computer, or machine interpretation, which requires the simultaneous interpretation and translation of the major languages. With Android as a platform, the proposed methodology creates a language translation application for Android devices that leverages firebase ML to handle problems in the most generally spoken languages. The Android-based program offers a solution for those who are unable to learn a language because they do not speak a common language, as well as for a variety of other reasons. This work applies the erudition

process to a previously unexplored sort of pre-translated language, in which a number of different designations have been made such as comparative corpus, equivalent text, bi-text, or multi-texts. This will assist in defining the upper limit of human translation in terms of cost, the number of languages available for translation, and efficiency.

2 Literature Review

The languages of the world are, without a doubt, diverse and many, and their distribution throughout the globe is constantly shifting and unexpected. There have already been 6,170 living languages, excluding dialects, identified, and characterized [1, 2]. In comparison with a human translator, an automated computer translator is incapable of performing tasks like taking into account grammar, idioms, and conventions or maintaining a meaning as close to that of the original as possible when translating into the target language [3, 4]. Translation of natural language text or speech by machine is used to solve issues of language impenetrability [5–7]. Machine translation is a technique for dealing with language impenetrability difficulties. Many advancements have occurred since the first machine translation software was developed to meet the demands of those who need precise and accurate communication. Unlike machine translation, machine learning finds its application in various fields [8–10]. This section provides the literature review of various research work in language translation.

It was published in a journal on translation that the article Android platform for machine translation—A focus on Yorùbá Language, which was produced on a mobile platform for more convenience, ease, and handiness, was published [11]. As far as decision support and data analysis are concerned, rough set theory is the mathematical instrument of choice [12]. It is necessary to make comparisons between a query, which is a word or phrase that needs to be solved, and the freshly produced corpus, and this is done using rough set theory.

For multiple language voice translator for English document, the use of Android was suggested as a possible option [13]. Specifically, the purpose of this article is to describe the design an Android framework whose purpose is to provide a solution to the language barrier by combining text for translation to voice, which can be done in multiple languages, in order to overcome the language barrier. A voice conversion software package was developed as part of the Android architecture used for the conversion of English-language text to vocal output in various native languages.

The translation tool [14], that is able to recognize the Arabic sign language gestures, has been designed to assist hearing-impaired people in integrating into society more easily. The purpose of this research is to develop a theoretical foundation for a smart translator to use in order to comprehend the Arabic sign language autonomous and fluid motions.

An Android application was developed using the help of the Tesseract OCR engine, the Bing translator, and the phones in built conversation feature [15], which

was then distributed to users. The suggested application deliverable has been evaluated by a variety of target end-users from a variety of dialectal contexts, and it has been determined that the usage benefits a number of operators in the field.

In order to translate American sign language into general text that can be used anywhere, a framework for Android-based devices has been developed [16] and released to the public. Skin segmentation is accomplished by the use of YCbCr systems, which are captured by the mobile camera. In order to recognize the sign, features are extracted from the picture using HOG and a set of features. The categorization process was done with the help of the support vector machine (SVM).

In a study on sign language translator software, OpenCV was employed [17]. Specifically, the development of a sign language translation application based on the Android-based OpenCV framework is the focus of this research; the color discrepancy is the driving force behind this technology. When it comes to anticipating the objective, the author often employs machine learning techniques.

Evelyn et al. [18] created an Android-based natural language processing system for translating from English to Igbo language translation. It was accomplished in Microsoft Hub to design and develop the Word document, reference system, and decoder; the system for training related documents and language transformation was developed in the Android environment and accessible via Android apps installed on phones.

Hakkun et al. [19] created an Android-based software that was capable of accurately translating sign language conveyed in written language by a deaf speaker into written language. This technique begins with the hand recognition using and conversion of the hand signals, which are both conducted by the OpenCV algorithm. The demonstration features were implemented in this application in order to educate users how to utilize sign language in a more intense manner.

Hidayatullah et al. [20] created a competence test for Android devices. Advancement was achieved via the use of a quiz format in this study. Various questionnaire approaches from the five kinds of quizzes (multiple-choice, true/false, matching chart, completion exam, and fill in the blank) were used in the study's solitary setting. This research's input data instruments take the form of components and problems arising from a competence examination.

Ogundokun et al. [21] have created an Android application that permits communication between the people who are either deaf or dumb. A regular person who does not understand the signals' etymological meaning. Half-duplex style communication software that serves as an effective mediator between deaf and dumb individuals and the general public by converting their hand signals into an audio medium that the general public can readily perceive, and vice versa, is available.

Ambar et al. [22] developed an experimental idea that was a sign language translation system powered by dual-sensor technology. In order to understand the movements of each finger and neck, the technology was implemented using a glove-based instrument that used two different types of radars, an accelerometer and five different components of flex radars.

There are several studies on local languages [18, 23], but there have been few attempts toward the construction of a language translator for android devices, particularly for multilingual users. This served as inspiration for us in the course of our current job.

3 Results

The proposed work translates the text or speech given as input in different languages to some specific language as selected by the user. As illustrated in Fig. 1, the user selects a particular language from the "from" spinner that contains 11 languages (English, Afrikaans, Arabic, Belarusian, Bulgarian, Bengali, Catalan, Czech, Hindi, Urdu, and Welsh). A specific language is selected in "to" spinner in which user wants to translate the written text or spoken text.

Fig. 1 From and to spinner

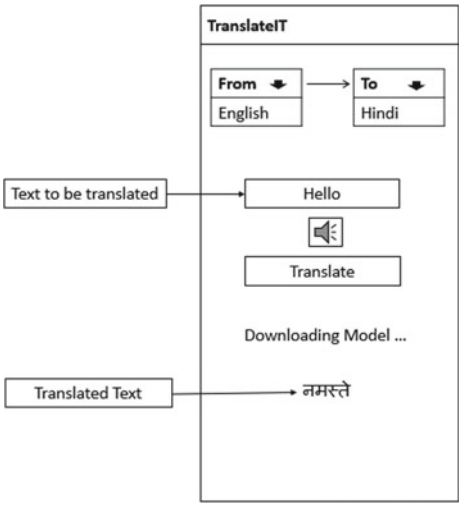


Fig. 2 English to Hindi translation

3.1 Results of Text to Text

This section describes the translation of written text into specific language selected by user in "to" spinner. The user writes the text in input text box using keyboard of mobile application. Then user clicks the "translate" button in application and output will be displayed in text format in the language chosen by user. Figure 2 illustrates the translation of English text into Hindi text. Likewise, we can convert any text written in language selected from "from" spinner to language selected in "to" spinner.

3.2 Results of Voice to Text

This section describes the translation of speech into specific language selected by user in "to" spinner. The user gives the input in speech form using either by clicking "press to speak" button or by using google voice typing option from the keyboard. Figure 3 illustrates the results of language translation by using google voice input.

4 Conclusion

The main purpose of the work presented is to create an android-based application to manage the communication gap present among tourists, language barrier among students learning online or offline and to invoke interest among people to learn new languages. Many more advancements and improvements can be made in the application to improve user interaction and make it more user-friendly. In the future,

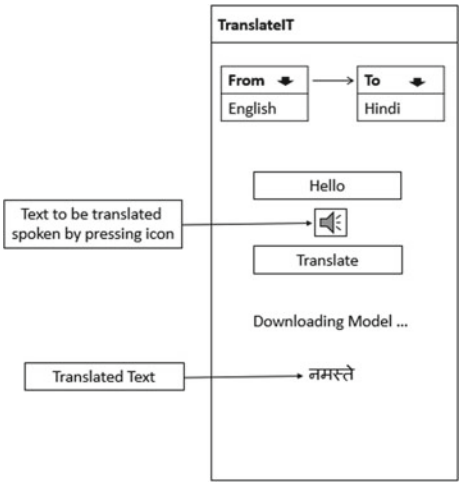


Fig. 3 Language translation using google voice input

many more languages can be added to the application to expand the usage of this application globally. Future, plans also include, but not limited to eye-tracking which can translate words which the user has his/her eyes on using the front camera of the mobile. Many other features like user-feedback, read-aloud, and pronunciation of translated words can also be added.

References

1. Mackey WF (1991) Language diversity, language policy and the sovereign state. Hist Eur Ideas 13:51–61. https://doi.org/10.1016/0191-6599(91)90113-D
2. National Reasearch Council (2011) Allocating federal funds for state programs for English language learners. National Academies Press, Washington, D.C. https://doi.org/10.17226/13090
3. Trappey AJC, Trappey CV, Wu J-L, Wang JWC (2020) Intelligent compilation of patent summaries using machine learning and natural language processing techniques. Adv Eng Inform 43:101027. https://doi.org/10.1016/j.aei.2019.101027
4. Saquete E, Tomás D, Moreda P, Martínez-Barco P, Palomar M (2020) Fighting post-truth using natural language processing: a review and open challenges. Expert Syst Appl 141:112943. https://doi.org/10.1016/j.eswa.2019.112943
5. Pathak A, Pakray P, Bentham J (2018) English–Mizo machine translation using neural and statistical approaches. Neural Comput Appl 1–17. https://doi.org/10.1007/s00521-018-3601-3
6. Ali MA (2020) Quality and machine translation: an evaluation of online machine translation of English into Arabic texts. OJML 10:524–548. https://doi.org/10.4236/ojml.2020.105030
7. Abazyan S, Mamikonyan N, Janpoladov V (2020) Interlanguage translation utility with integrated machine learning algorithms. OA Lib 07:1–5. https://doi.org/10.4236/oalib.1106318
8. Chaudhary D, Kumar S, Dhaka VS (2021) Estimating crowd size for public place surveillance using deep learning. In: Ahmed KR, Hassanien AE (eds) Deep learning and big data for intelligent transportation: enabling technologies and future trends. pp 175–197. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-65661-4_9

9. Chaudhary D, Kumar S, Dhaka VS (2021) Video based human crowd analysis using machine learning: a survey. Comput Methods Biomech Biomed Eng: Imaging Visual 1–19. https://doi.org/10.1080/21681163.2021.1986859
10. Sharma VP, Gaded AS, Chaudhary D, Kumar S, Sharma S (2021) Emotion-based music recommendation system. In: 2021 9th international conference on reliability, infocom technologies and optimization (trends and future directions) (ICRITO). IEEE, pp 1–5. https://doi.org/10.1109/ICRITO51393.2021.9596276
11. Olaide F, Kayode AB, Adewale OS, Adetunmbi AO (2018) Android platform for ma chine translation-a focus on Yorùbá Language
12. Zhang C, Li D, Kang X, Song D, Sangaiah AK, Broumi S (2020) Neutrosophic fusion of rough set theory: an overview. Comput Ind 115:103117. https://doi.org/10.1016/j.compind.2019.07.007
13. Hanumante V, Debnath R, Bhattacharjee D, Tripathi D, Roy S (2014) English text to multilingual speech translator using android
14. Ahmed AM, Abo Alez R, Tharwat G, Taha M, Belgacem B, Al Moustafa AMJ (2020) Arabic sign language intelligent translator. Imaging Sci J 68:11–23. https://doi.org/10.1080/13682199.2020.1724438
15. Ramiah S, Liong TY, Jayabalan M (2015) Detecting text based image with optical character recognition for English translation and speech using Android. In: 2015 IEEE student conference on research and development (SCOReD). IEEE, pp 272–277. https://doi.org/10.1109/SCORED.2015.7449339
16. Lahoti S, Kayal S, Kumbhare S, Suradkar I, Pawar V (2018) Android based American Sign Language recognition system with skin segmentation and SVM. In: 2018 9th International conference on computing, communication and networking technologies (ICCCNT). IEEE, pp 1–6. https://doi.org/10.1109/ICCCNT.2018.8493838
17. Triyono L, Pratisto EH, Bawono SAT, Purnomo FA, Yudhanto Y, Raharjo B (2018) Sign language translator application using openCV. IOP Conf Ser: Mater Sci Eng 333:012109. https://doi.org/10.1088/1757-899X/333/1/012109
18. Evelyn CC, Bennett EO, Taylor OE (2019) A natural language processing system for English to IGBO language translation in android
19. Hakkun RY, Baharuddin A (2015) Sign language learning based on Android for deaf and speech impaired people. In: 2015 International electronics symposium (IES). IEEE, pp 114–117. https://doi.org/10.1109/ELECSYM.2015.7380825
20. Hidayatullah MS, Fahrizal R (2020) UKOPAI, android-based development of comptence test for Arabic–Indonesian translators. In: Proceedings of the 2nd international conference on Islam, science and technology (ICONIST 2019). Atlantis Press, Paris, France. https://doi.org/10.2991/assehr.k.200220.040
21. Ogundokun RO, Awotunde JB, Misra S, Segun-Owolabi T, Adeniyi EA, Jaglan V (2021) An android based language translator application. J Phys: Conf Ser 1767:012032. https://doi.org/10.1088/1742-6596/1767/1/012032
22. Ambar R, Fai CK, Choon CC, Wahab MHA, Jamil MMA, Ma'Radzi AA (2018) Preliminary design of a dual-sensor based sign language translator device. In: Ghazali R, Deris MM, Nawi NM, Abawajy JH (eds) Recent advances on soft computing and data mining. Springer International Publishing, Cham, pp 353–362. https://doi.org/10.1007/978-3-319-72550-5_34
23. Adubi SA, Misra S (2016) Syllable-based text compression: a language case study. Arab J Sci Eng 41:3089–3097. https://doi.org/10.1007/s13369-016-2070-1

Visualizing Commenters Opinions Through Topic Analysis

Ayush Soni, Akhtar Rasool, Aditya Dubey, and Nilay Khare

Abstract With the rapid development in the computer science and technology domain, the eruption of available data is observed in this decade. Online platforms are becoming more and more capable day by day and therefore can capture thousands of customer reviews and comments on a single post. Dealing with this gigantic amount of high-dimensional text data leads to several problems for the post owner and data analysts. Furthermore, a significant percentage of this high-dimensional text data is not important and can be proficiently concentrated to lower dimensions by using several advanced dimensionality reduction methods. Topic modeling methods are used to summarize text data efficiently and are a good way of analyzing a huge amount of text data. In the recent years, many topic modeling approaches are introduced and are used to gain fruitful insights from the considered dataset. This paper aims to suggest the most effective known model by analytically comparing several existing topic models on the taken dataset. Also, some modifications to the existing algorithms are suggested to generate more understandable and accurate results.

Keywords Topic modeling · Dimensionality reduction · Text mining · Feature extraction · Machine learning · YouTube · Latent Dirichlet allocation · Text analysis · Visualization

1 Introduction

In the last decade, a huge advancement in online platforms is observed along with the swift development of high-end technologies [1]. For example, in the social media field, tremendous advancement is observed and platforms like Twitter, Facebook, YouTube, Instagram, and WhatsApp have become a necessary need for everyone. Also with the cheap and high-speed Internet services, many users are using all these platforms every day which leads to the generation of a huge amount of data. With the fast expansion of networking and data storage capacity, these platforms provide users

A. Soni (✉) · A. Rasool · A. Dubey · N. Khare
Maulana Azad National Institute of Technology, Bhopal, India
e-mail: ayush17soni@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_27

with interactive features like commenting, giving star ratings, and posting pictures, etc., which leads to generation of enormous text data. Therefore, summarization of digital data is becoming a key field of research in the science and engineering domain [2]. The substantial size of text data causes several challenges, such as problems in processing the document, problems in extracting relevant information, problem in data analysis, data heterogeneity, and high computation time [3]. Therefore, to analyze text data successfully, major efforts are required [4].

In the text world, dimensions are the words of vocabulary and text documents are represented as a vector of words counts over a vector-space model of documents, where the dimensionality of the documents can be over 15,000. Also, high-dimensional text data is implicitly sparse [5]. For example, a word with a frequency of 50 in a particular document might not appear in any other document [6]. To overcome these problems, exploration of text dimension reduction is important [7]. Dimension reduction, clustering, and topic modeling have interesting relationships between them [8]. As for text mining, these methods illustrate documents in a way that discloses their internal structure [9]. Essentially, topic modeling combines soft clustering with dimension reduction, as there are documents that are a combination of hidden topics, which is similar to both document clusters and compact representation of a corpus [10]. In topic modeling, original features are important for defining the topics and detecting the topics present in each document, which results in an understandable representation of documents and is useful for analyzing the theme of the document. Traditional topic modeling approaches include latent semantic analysis (LSA) and latent Dirichlet allocation (LDA). Both of these methods take a bag of words in matrix format as input, where LSA focuses on reducing the dimension of this matrix, while LDA solves topic modeling problems [11].

This work explores some text analysis and machine learning algorithms and aims to present a comparative analysis of the existing topic modeling approaches over the dataset taken from youtube.com and proposes an enhancement that results in a better analysis of the taken dataset. The organization of this paper is as follows: Sect. 2 will discuss the essential background details of the text analysis and techniques used. Section 3 covers an explanation of the methodology used. Section 4 discusses the results and provides an analytical comparison using enhanced performance metrics. Finally, Sect. 5 presents the concluding remarks.

2 Background

2.1 Non-Negative Matrix Factorization (NMF)

NMF is one of the most useful algorithms for feature extraction in unstructured text mining [12]. It is very useful when many attributes are there and attributes have weak predictability. Each attribute extracted by NMF is a linear combination of the original features. In NMF, input matrix I of dimension $m \times n$ gets factorized into

two matrices $M1$ and $M2$ of dimension $m \times k$ and $n \times k$, respectively, where I is essentially the document term matrix in text dataset and k is the reduced dimensions. NMF focuses on the problem of minimizing $I - M_1 M_2^2$ with respect to $M1$ and $M2$, such that $M1$ and $M2$ are non-negative. There are many methods to find the $M1$ and $M2$, among which Lee and Seung's method is the most popular because of its simplicity of implementation [13]. It starts with initializing $M1$ and $M2$ with non-negative values, and updation of $M1$ and $M2$ is done iteratively to make their product approximately equal to I.

2.2 LSA

To capture latent topics, LSA tries to find the context around the words, i.e., words are similar if they occur in a similar context [14]. In this, a document is treated as a bag of words, therefore, semantic and syntactic information are ignored. Usually, it uses the TF-IDF matrix which contains scores for each word in the document and tries to reduce it to k topics or dimensions [15]. Dimension reduction is done by making use of singular value decomposition (SVD).

$$A = USVT \tag{1}$$

In (1), A is SVD matrix, U is a representation of documents in vector form, V is a representation of terms in vector form, S is a diagonal matrix, and T is a hyperparameter for the number of topics. With the help of cosine similarity, similar topics and documents can be found using this SVD matrix. LSA often requires a large set of documents to yield accurate results.

2.3 LDA

LDA is considered as one of the most effective unsupervised topic modeling approaches [16]. LDA tends to find the hidden/latent topics within a corpus by assuming that documents are the distribution of topics and topics are the distribution of words. Figure 1 shows a generative model assumed by LDA, where all the shaded circles are observed variables while not shaded are unobserved, α is the Dirichlet parameter, θ_{d} is the topic proportion of each document, $Z_{\mathrm{d,n}}$ is topic assignment for each word, $W_{\mathrm{d,n}}$ is the word that we observe in documents, β is topic distribution for a corpus, which is the distribution of words, η is topic hyperparameter, N is the number of words per document, D is the number of documents in the corpus, and K is the total number of topics.

Many extensions of LDA are proposed in the recent years. Blei and Lafferty introduced the dynamic topic model (DTM) which can be used to obtain the evolution of topics when a sequentially arranged corpus of documents is used [17]. Labeled

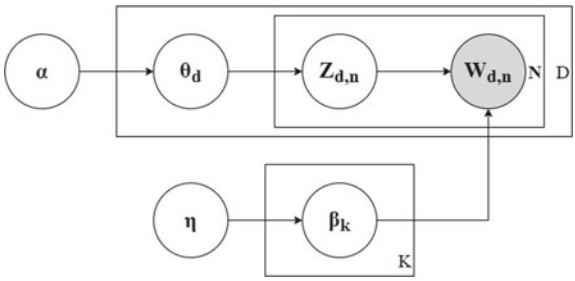


Fig. 1 Classical LDA generative model

latent Dirichlet allocation (LLDA) is also introduced which works on assumption that labels are already there for each document [18]. Another approach to detect different languages present in corpus is introduced [19]. Conceptual dynamic latent Dirichlet allocation (CDLDA) is one of the recent advancements which is focused on the analysis of communicational conversation [20].

2.4 Topic Coherence Measure

Judging the appropriateness of discovered topics from topic modeling algorithms has always been a problematic task as it required human intelligence and analysis. Major research is carried out to quantification of the quality of generated topics. Measure like the predictive likelihood of data is used for this purpose [21]. However, such a measure does not correlate positively with human interpretability. And it is quite important when produced topics are used to understand certain trends. Hence, researchers proposed a qualitative measure known as topic coherence measure to automatically understand the coherence of topics.

3 Methodology

3.1 Dataset Selection, Text Preprocessing and Selection of Number of Topics

Dataset for this work is scraped from youtube.com which contains columns like user_id, number_of_likes, and comment_text of over 10,000 + users, posted below the BBC news report on Ukraine and Russia conflict. As this paper is concerned with the analysis of the theme of the comments, all other unnecessary columns are removed and only comments text is taken into account [22]. Also, punctuation removal, stop words removal, and lemmatization are performed before using the dataset in any machine learning algorithm. Preprocessing of raw data is always

needed before any textual analysis; the Natural Language Toolkit in Python has some useful functionalities to so and is therefore used here [23]. As this work is concerned with the effective analysis of the taken dataset, this work compares several well-known topic modeling algorithms like NMF, LSA, and LDA. As number of topics is a user-defined parameter in most topic modeling approaches. It is important that number of topics is chosen with caution as, a small topic number may combine two closely related topics, while a big number can lead to generation of topics that are hard to interpret. To select the best number of topics, LDA tuning is done through Cv coherence measure. The best Cv value observed during this process is, when the number of topics is selected as 4.

3.2 Iterative Latent Dirichlet Allocation (iLDA) and Iterative Non-negative Matrix Factorization (iNMF)

This paper proposes an extension of the classical LDA and NMF topic modeling method in order to do a better textual analysis of the commenter dataset. Figure 2 shows a flowchart of the proposed iLDA and iNMF techniques. It starts with the classical LDA or NMF model, where input is the corpus of documents after getting topics from LDA or NMF each word distribution within a topic is observed and the threshold is defined. This threshold helps to remove unnecessary words which affect the LDA or NMF algorithm negatively. After removal, new dictionary is used to train LDA or NMF model again.

Pseudocode of the iterative algorithm used with LDA and NMF is given shown in Fig. 3. Initialization is done with the classical LDA in iLDA, while NMF is used with iNMF. Gensim library of Python is used for classical implementation. The value of preset helps here to reach better results in fewer iteration and allow us to filter out unnecessary words in less than 4 iterations. It is observed that the proposed iterative algorithm is achieving a better coherence score in 2nd iteration for iLDA and 3rd iteration for iNMF. Obtained coherence score of iLDA is 0.4227, and for iNMF, it is 0.7224 which is better than older topic modeling algorithms. Therefore, this work

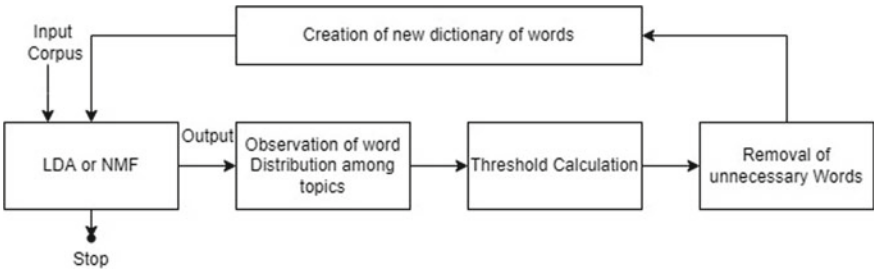


Fig. 2 Flowchart of iLDA and iNMF

```
Pseudocode of iLDA and iNMF

Essential preprocessing steps as defined
Conversion of dataset into tokens
Tn = Number of topics //Selected by tuning a classical LDA model
Training of {LDA or NMF}
Initailly previous_model is { LDA or NMF}
preset = 0.0005 //preset weight is used to reduce number of iterations
for(iterations = 0 to 3)
{
    Loading previous {LDA or NMF} model
    Topics = output of previous {LDA or NMF} model
    reducedset = 0  //Reduced set of words
    for(i = 0 to Tn )
    {
        wordinfo = "words*weight" //"word"*0.016
        topicweight = 0 //Total weight of all the words within a topic
        topicwords = 0 // List of all the words of a topic
        listofwordweights = 0 // List of all the words weights of a topic
        for(j = 0 to countof(wordinfo))
        {
            wordvalue = word part in wordinfo
            wordweight = weight part in wordinfo
            if wordweight > 0 // Remove all the words whose weights are 0
            {
                topicwords = topicwords.append(wordvalue)
                listofwordweights = listofwordweights.append(wordweight)
                topicweight = topicweight + wordweight  // Total topic weight calculation
            }
        }
        theshold = topicweight/countof(wordweights) + preset  //Setting threshold value
        for(k = 0 to countof(topicwords))
        {
                if(listofwordweights[k]>=threshold)
                        reducedset.append(listofwordweights[k]) //Add words to reducedset of words
        }
    }
    preset = preset*2
    newtokenlist = reducedset
    Train new {LDA or NMF} Model with reducedset
    Save new {LDA or NMF} for next iteration
}
```

Fig. 3 Pseudocode of iLDA and iNMF

uses the iLDA and iNMF techniques to do analysis and compares it with traditional topic modeling algorithms.

4 Results and Visualization

All of the above-used topics modeling algorithms are implemented in Jupyter Notebook using Python programming language, and the dataset used is scraped from youtube.com. As this paper aims at exploring important aspects of the taken dataset, it proposes two extensions of LDA and NMF and tries to figure out which topics are more talked about in the comment section and intern helps the community to discover the important topics people are talking about in the comment section which indeed save the time of going through thousands of comments. The word count and importance of topic keywords detected by classical LDA are not very well-formed and cannot be judged easily. Figure 4 shows word count and importance of topic

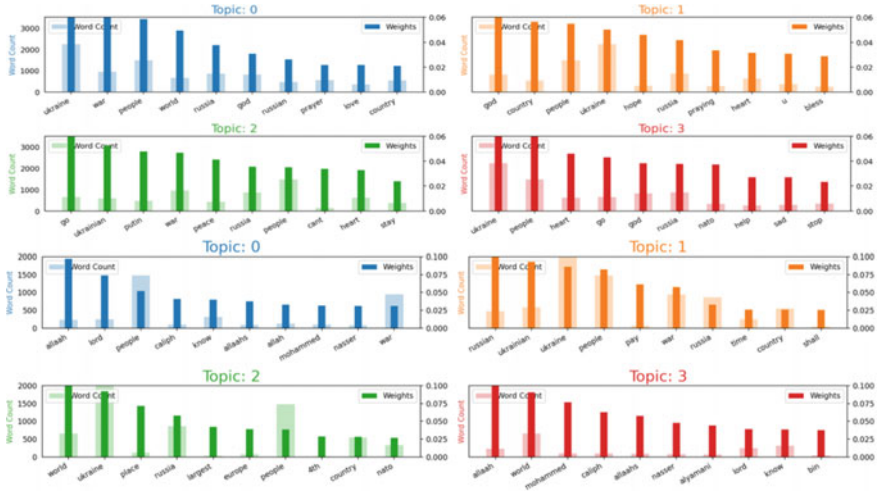


Fig. 4 Word count and importance of topic keyword for iLDA (top) and iNMF (bottom)

keywords detected by iLDA and iNMF, and during analysis, these topics are much more interpretable and can be helpful in effective analysis of any dataset. Table 1 gives the top 10 keywords detected by iLDA and iNMF where, "Russia and Ukraine War," "Prayers for Ukrainian people," "Appeal for peace," and "NATO help" are the discovered topics by iLDA. While "Prayers for people," "Russia and Ukraine War," "European Conflict," and "Prayers" are the discovered topics by iNMF.

5 Conclusion

User comments/reviews are an invaluable resource for analysts as it depicts the likes and dislikes of users and their mindset about the discussed post/article. As textual data is very high-dimensional and hard to analyze, this work tries to reduce the dimensions as to extract useful information hidden inside the user's free-text comments. For effective analysis, LDA, NMF, and LSA topic modeling algorithms are compared based on the topics generated by them and obtained coherence score. Also, an extension of LDA and NMF known as iLDA and iNMF respectively are proposed, which uses results of classical LDA and NMF to reach more accurate results through iterations. Results show that the proposed iLDA approach has a better coherence score for the taken dataset in the 2nd iteration while iNMF is giving better coherence in the 3rd iteration hence these models are used to visualize the comments of the users. It is observed that the proposed method is useful to generate more interpretable topics which are easy to analyze. For future work, this approach can be analyzed in many other datasets and can be compared to the existing methods.

Table 1 Discovered topics by iLDA and iNMF

S. No	Technique and topic number	Top 10 words discovered	Topic discovered
1	iLDA Topic 0	Ukraine, War, People, World, Russia, God, Russian, Prayer, Love, Country	Russia and Ukraine War
2	iLDA Topic 1	God, Country, People, Ukraine, Hope, Russia, Praying, Heart, You, Bless	Prayers for Ukrainian people
3	iLDA Topic 2	Go, Ukrainian, Putin, War, Peace, Russia, People, Cant, Heart, Stay	Appeal for peace
4	iLDA Topic 3	Ukraine, People, Heart, Go, God, Russia, NATO, help, Sad, Stop	NATO help
5	iNMF Topic 0	Allah, Lord, People, Caliph, Know, Allaahs, Allaah, Mohammed, Nasser, War	Prayers for people
6	iNMF Topic 1	Russian, Ukrainian, Ukraine, People, Pay, War, Russia, Time, Country, Shall	Russia and Ukraine War
7	iNMF Topic 2	World, Ukraine, Place, Russia, Largest, Europe, People, 4th, Country, NATO	European Conflict
8	iNMF Topic 3	Allaah, World, Mohammed, Caliph, Allaahs, Nasser, Alyamani, Lord, Know, Bin	Prayers

Also this iterative approach can be applied to other topic modeling algorithms and can be analyzed for better results.

References

1. Dubey A, Rasool A (2021) Efficient technique of microarray missing data imputation using clustering and weighted nearest neighbour. Sci Rep 11:1–12
2. Sharma MK, Kumar P, Rasool A, Dubey A, Mahto VK (2021) Classification of actual and fake news in pandemic. In: Fifth international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), pp 1168–1174
3. Vyas P, Sharma F, Rasool A, Dubey A (2021) Supervised multimodal emotion analysis of violence on doctors tweets. In: International conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC), pp 962–967
4. Soni A, Rasool A, Dubey A, Khare N (2022) Data mining based dimensionality reduction techniques. In: International conference for advancement in technology (ICONAT), pp 1–8
5. Espadoto M, Martins RM, Kerren A, Hirata NST, Telea AC (2021) Toward a quantitative survey of dimension reduction techniques. IEEE Trans Visual Comput Graph 27(3):2153–2173
6. Abdulwahab HM, Ajitha S, Saif MAN (2022) Feature selection techniques in the context of big data: taxonomy and analysis. Appl Intell

7. Anowar F, Sadaoui S, Selim B (2021) Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). Comput Sci Rev 40:1–13
8. Luan Y, Eisenstein J, Toutanova K, Collins M (2021) Sparse, dense, and attentional representations for text retrieval. Trans Assoc Comput Linguist 9:329–345
9. Kherwa P, Bansal P (2020) Topic modeling: a comprehensive review, EAI Endorsed Trans Scalable Inf Syst 7(24)
10. Yau CK, Porter A, Newman N, Suominen A (2014) Clustering scientific documents with topic modeling. Scientometrics 100:767–786
11. Mohammed SH, Al-augby S (2020) LSA & LDA topic modeling classification: comparison study on e-books. Indonesian J Electr Eng Comput Sci 19(1):353–362
12. King G, Patrick L, Roberts ME (2014) Computer-assisted keyword and document set discovery from unstructured text. Am J Polit Sci 61(4):971–988
13. Lee DD, Seung HS (2000) Algorithms for non-negative matrix factorization. Adv Neural Inf Process Syst 13:556–562
14. Wiemer-Hastings P, Wiemer-Hastings K, Graesser A (2004) Latent semantic analysis. In: International joint conference on artificial intelligence
15. Aizawaand A (2002) An information-theoretic perspective of TF–IDF measures. Inf Process Manage 39:45–65
16. Dubey A, Rasool A (2020) Clustering-based hybrid approach for multivariate missing data imputation. Int J Adv Comp Sci Appl 11(11):710–714
17. Blei DM, Lafferty JD (2006) Dynamic topic models. In: 23rd international conference on machine learning. ACM
18. Ramage D, Hall D, Nallapati R, Manning CD (2009) Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In: Conference on empirical methods in natural language processing, vol 1. pp 248–256
19. Lui M, Lau JH, Baldwin T (2014) Automatic detection and language identification of multilingual documents. Trans Assoc Comput Linguist 2:27–40
20. Yeh JF, Tan YS, Lee CH (2016) Topic detection and tracking for conversational content by using conceptual dynamic latent Dirichlet allocation. Neurocomputing 216:310–318
21. Dubey A, Rasool A (2022) Recent advances and applications of deep learning technique. J Appl Inf Tech 100(13):500–509
22. Cunningham-Nelson S, Baktashmotlagh M, Boles W (2019) Visualizing student opinion through text analysis. IEEE Trans Educ 62(4):305–311
23. Bird S, Klein E, Loper E (2009) Natural language processing with python: analyzing text with the natural language toolkit. O'Reilly Media, Beijing, China

A Survey of Different Modulation Schemes and Channel Modeling Techniques of a VLC System

Supratim Subhra Das and Md. Asraful Sekh

Abstract This paper presents a brief survey on different modulation schemes of a visual light communication (VLC) system, based on which the performance of different modulation schemes has been compared.

Keywords VLC · Flickering · Dimming · PWM · PPM · OOK

1 Introduction

Visible light communication (VLC) system is becoming an apparent research field in today's world. The main advantage of optical communication over RF transmission is high data transmission rate. With respect to some other wireless communication technologies (e.g., Wi-Fi, LTE), VLC has got few potential features like higher bandwidth, higher reliability and a more secured communication. Unlike the radio frequency (RF) transmission, the optical communication technology is likely to be more immune to the problem of security, since the optical signal cannot be passed through the walls. LED or LASER is common optical sources in optical communication technology, with wavelengths ranging from 380 to 750 nm and spectral frequencies ranging from 430 to 790 Hz. Pulse position modulation, OOK modulation, pulse amplitude modulation, OFDM modulation, pulse width modulation and NRZ modulation are some important modulation schemes and channel modeling techniques used in VLC using single and multicarrier. The performance of several modulation methods for VLC is compared in this paper by looking at some key characteristics relevant to the elemental behavior of the VLC system. These are the light source's spectral efficiency, bandwidth efficiency, power and dimming factor. More VLC technology criteria are used to evaluate the communication system's efficiency and reliability, such as data rate, SNR and BER. The benefits of visible light communication include a wide spectrum, efficiency and the most recent advancements in LED technology with faster switching times, all of which have led researchers to

S. S. Das (✉) · Md. A. Sekh
Aliah University, ll-A/27Action Area II, Newtown Kolkata, West Bengal 700160, India
e-mail: supratim2014das@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_28

conduct additional research in this area. Conventional radio frequencies below 6 GHz are incapable of delivering higher data rates. Traditional radio frequency cannot attain higher data rates below 6 GHz, however, VLC can typically reach 300 THz of bandwidth with higher data rates across a short communication range. In order to obtain higher data rate, optical sources (e.g., LEDs) are implemented in such multiple-input multiple-output (MIMO) approach [1] in VLC system. To achieve idea data rates for transmission, improved flickering and dimming sustainability are two key phenomena associated to visible light spectrum. Flickering is a term for the variation in light intensity that can have a noticeable and harmful impact on human physiological changes. There is a solution to restrict light intensity changes by using a maximum flickering time period to mitigate this problem (MFTP). The ideal fluctuation number is regarded safe and sound for frequencies more than 200 Hz (MFTP5) [2]. As a result, the modulation mechanism in the VLC system is incapable of generating any variation within or between data frames. Moreover, the dimming sustainability in VLC enables power and energy saving with increase in its efficiency. The low light levels of dim light sources make human eyes to react by amplifying the pupil, enabling the eyes to receive more light. The dim light sources are hence more desirable for communication. As illustrated in Fig. 1, this paper discusses some important modulation strategies for VLC.

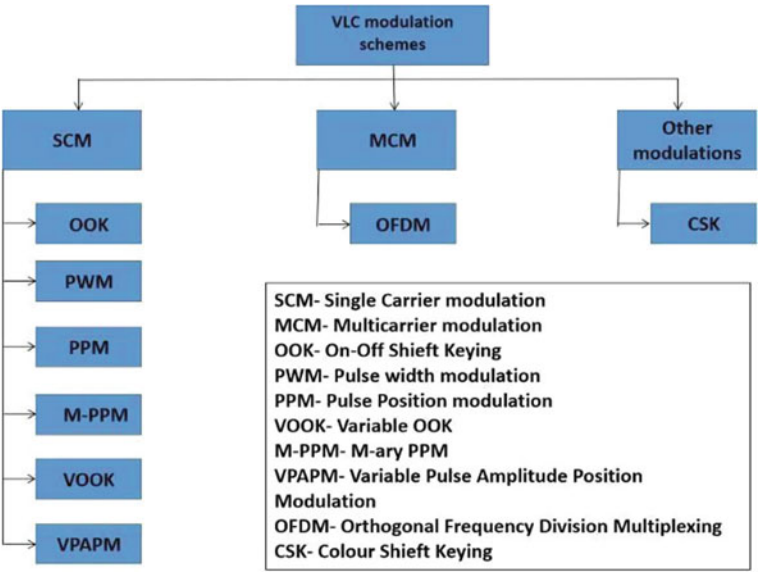


Fig. 1 Different modulation schemes for VLC

2 Modulation Techniques

VLC system usually adapts various modulation schemes. An important difference between RF and VLC technology is that the signal cannot be encoded over amplitude and phase in VLC. The light intensity is modulated and demodulated using direct detection demodulation. The choice of optical sources in a VLC system depends on the following features of light source.

- *Dimming*

  Different amounts of light intensity are required to conduct different types of actions, according to [3]. The illuminance should be in the range of 30–100 Lux to reduce the darkness and conduct the visual operation or work. For home and office uses, however, light intensities between 300 and 1000 Lux are sufficient. With the progress of LED driver circuits, it is becoming increasingly important to comprehend the influence of a random amount of LED light intensity on human vision. Figure 2 depicts the nonlinear response [4] analyzed between perceived and measured beam.

  The calculation of the perceived light from the quantified light is shown in Eq. (1).

$$\text{Perceived Beam } (\%) = 100\sqrt{((\text{Quantified Beam})/100)} \tag{1}$$

  The human eye perceives the measured light in a different way due to the contraction/enlargement of the pupil. It is also reported that if the lamp source is darkened 1%, the human eye perceives a 9% increase of measured light with a 10% increase in overall light intensity. The users can select any arbitrary level of light intensity based on the required energy saving and application, with no impact of dimming in the communication. The modulation is applied on the signal that receives information from a dim source of light without affecting the data transmission.
  Flickering

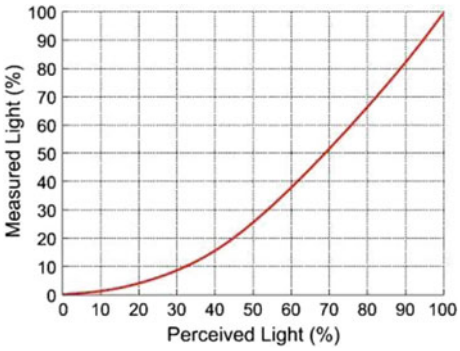


Fig. 2 Perception of actual measured light by human eye

It is preferable for VLC modulation methods not to allow the human eye to notice fluctuations in light intensity, as they might induce major physiological changes in humans [5]. As a result, light intensity variations must occur faster than the human eye's perception rate. To avoid any harmful effects, the flickering of the light intensity should be made quicker than 200 Hz, according to the IEEE 802.15.7 standard [6]. It was also claimed that before supplying larger data rates for transmission, modulation techniques for VLC should alleviate the problem of oscillations. The flickering is mostly produced by the extended sequence of 0s and 1s, which slows down the rate of change of light intensity, resulting in fluctuation. The concept of using Run Length Limited (RLL) codes can alleviate the long run of 0s and 1s in the output sequence, ensuring balanced repetitions of 0s and 1s. RLL codes such as Manchester, 8B10B and 4B5B are widely used. In Manchester coding, the "down" transition ("10") is represented by "0," and the "up" transition ("01") is represented by "1." To adjust the duplication, 4B6B and 8B10B are used to map 4 bits symbols to 6-bit symbols and 8-bits symbols to 10-bit symbols, respectively. Manchester coding's extra data bit requirements make it suited for low data rate services and greater management of the 0s and 1s balance. The modulation methods used in visible light communication (VLC) system consist of (1) OOK (2) PM (3) OFDM and (4) color shift keying (CSK) modulation. The next section will look at how the modulation techniques outlined above help VLC communicate with the dimming factor.

2.1 On–Off Shift Keying (OOK)

The data bits 1 and 0 are identified by the on and off of LEDs in OOK. The intensity of light is lowered in the off state, but it is not totally turned off. Because of its simplicity and ease of implementation, it is commonly used in wired communication. In the previous studies, the OOK modulation was mostly used for VLC using a white LED that produces a blue emitter with a yellow phosphor. White LED has a limited transfer speed due to the sluggish response time of yellow phosphor (few megahertz [7]). For a VLC with a data rate of 10 Mbps, the NRZ-OOK with the white LED is recommended. To compensate for the yellow component's slow response time, a blue filter is applied, resulting in a data rate of 40 Mbps [7]. The performance can be improved even further by selecting the right photodiode. The avalanche photodiode performs better at the receiver side than the P-I-N photodiode, as shown in [8]. The avalanche photodiode contributes to a data rate of 230 Mbps. Avalanche photodiode produces white light with a mix of RGB frequencies, has a fast response time, and requires three independent driving circuits to escape. IEEE standards [6] suggested the two ways as IEEE 802.15.7, which investigates dimming support OOK as a modulation scheme.

i. This technology redefines the ON and OFF levels by allocating different levels of light intensity to achieve the appropriate level of dimming without the use of an overhead bit. It maintains the same data rate as NRZ-OOK modulation, but the

communication range is limited at low dimming levels. The disadvantage of this technique is that it uses lower light intensities in LEDs for ON/OFF, resulting in low power driving circuits and changes in radiated shade [9].

ii. Compensation periods: In this strategy, additional compensation periods with the same ON and OFF modulation levels are employed to solve the problem, and the length of the periods is decided by the desired amount of dimming. The ON times are added when a dimming level of 50% is required; otherwise, the OFF periods are added. The authors [10] suggest a method for determining dimming intensity depends on the percentage of time spent transmitting active data inside the transmission period T.

The dimming level of D is obtained in [10] as

$$\gamma = \begin{cases} (2 - 2D) \times 100W : D < 0.5 \\ 2D \times 100 : D > 0.5 \end{cases} \tag{2}$$

OMPPM can improve MPPM's spectral efficiency without requiring an increase in data transmission capacity in a noiseless channel. The performance analysis of the additional noisy channel is exemplified in [11]. The performance of a noisy channel in OPPM is better if it has fewer time slots.

The maximal communication efficiency E_{D} can be estimated [10] as follows using information theoretic entropy after achieving the appropriate dimming level D with OOK:

$$E_D = D \log_2 D - (1 - D) \log_2 (1 - D) \tag{3}$$

According to Eq. (3), the dimming level's system efficiency achieved maximum proficiency at 50% of dimming. When the dimming intensity is between 0 and 100%, the efficiency declines linearly. The use of compensating periods in dimming is to blame for the drop-in data rate. To solve this problem, the authors of [12] proposed using inverse source coding at the required level of dimming to keep the data rate high.

2.2 Pulse Modulation Methods

The OOK modulation approach is limited to have lower data rates while preserving variable dimming levels. As a result, researchers chose several modulation strategies based on pulse position and width, which are summarized below.

2.2.1 Pulse Width Modulation (PWM)

Dimming can be achieved using PWM, which is an effective modulation approach. The pulse width is set to be balanced based on the needed level of dimming in this system. Based on the dimming requirement, the data transmission rate can be modified. The authors in [13] further notes that larger data rates can be achieved with this modulation approach with any level of dimming from 0 to 100%. PWM achieves dimming without reducing light intensity, which is one of its most important advantages; thus, no color shift is required for on and off levels in LEDs. PWM data rates are found to be limited to 4.8 kbps in [13].

2.2.2 Pulse Position Modulation (PPM)

Dimming can be achieved using PWM, which is an effective modulation approach. The width of pulses is set to be balanced based on the needed level of dimming in this system. Due to PPM's poor spectrum efficiency and data rate, an alternate pulse position-based modulation approach has been proposed over time. The overlap PPM (OPPM) approach allows for more symbols to be sent in a single pulse [14]. With a wide variety of dimming intensities, the authors of [15] proved that OPPM achieves higher spectral efficiency with higher data rates. Bai et al.[16] introduced another sort of PPM called multi-pulse PPM (MPPM). When compared to OPPM, MPPM has been found to have a greater spectral efficiency. In [17], a hybrid of OPPM and MPPM methods was presented, which was dubbed overlapping MPPM (OMPPM). A single optical signal is represented by many pulse positions in this technique. It has been stated that more pulses per symbol [18]. The differential PPM (DPPM) is a modulation method for VLC that was proposed in [19]. The OFF symbol has been deleted in DPPM, and the following symbol begins immediately after the previous symbol ends. For a given bandwidth, the DPPM consumes less power in the communication channel than the PPM. Zwillinger [20] also presented the differential overlapping PPM (DOPPM), which outperforms PPM, DPPM and OPPM in terms of spectral and cut-off performance. Another modulation approach termed EPPM, as demonstrated in [19], can solve the flickering problem. EPPM [20] changes the length of symbols and the number of pulses per symbol to allow fading. Multi-level EPPM (MEPPM), in contrast to EPPM, we can increase constellation size by adding multiple amplitudes level and spectrum efficiency [21]. The IEEE 802.15.7 standard proposes Variable PPM (VPPM), a pulse modulation technique in which the varying pulse positions are expressed using bits. PPM and PWM are combined in this modulation technique, which also allows for custom the levels of by varying the pulse width.

2.2.3 OFDM

Because of its nonlinear frequency response, any single carrier modulation scheme suffers from severe intersymbol interference (ISI). OFDM is a multicarrier modulation technology capable of effectively overcoming the conflict between intersymbol interference (ISI) and multipath fading. For VLC applications, OFDM was initially proposed in [22]. OFDM splits the bandwidth into numerous subcarriers and sends the modulated data in concurrent sub-streams. Although OFDM minimizes intersymbol interference, it is a difficult modulation technique to apply. Intensity modulation/direct detection (IM/DD) is a modulation technology that was initially used in a VLC system. By putting Hermitian symmetry constraints on the subcarriers, complex-valued bipolar signals can be transformed to real-valued unipolar signals in OFDM. The OFDM technique can be divided into two categories in this regard.

i. DC-biased optical OFDM (DCO-OFDM).
ii. Asymmetrically clipped optical OFDM (ACO-OFDM).

Only odd subcarriers are modulated in ACO-OFDM, resulting in a symmetric time domain signal [23]. DCO-OFDM technique is applied to modulate all subcarriers ([22–24]), and the signal is made unipolar by adding. According to [25], which analyzes both OFDM systems, the clipping distortion in LED is worse in DCO-OFDM than in ACO-OFDM. LED nonlinearity may impede OFDM-based VLC systems that produce higher peak-to-average power ratios, which is the OFDM VLC system's main flaw (PAPR). To overcome this problem, it was proposed in [26] to operate the LED in a small region, where the optical power and driving current are quasi-linear. Using only a single LED, OFDM can attain link rates in the hundreds of Mbps [27]. Color shift keying (CSK) is a technique that allows you to change the color of your image. The IEEE 802.15.7 standard [6] created the CSK modulation specifically for VLC to address the limited dimming support and lower data rate of existing modulation techniques. A solution was proposed in and [26] to solve this problem by operating the LED in a narrow region, where the optical power and driving current are quasi-linear. With only a single LED, OFDM may achieve link rates in the hundreds of gigabits per second [27]. (4) Color shift keying (CSK): The IEEE 802.15.7 standard [6] designed the CSK modulation specifically for VLC to tackle the restricted dimming support difficulties and lower data rate of other modulation techniques.

3 Comparison of Different Modulation Schemes

In this part, the performance of several modulation schemes such as MPPM, PPM, PWM, VPAPM, CSK, OFDM, DIPPM and OOK-NRZ modulation is compared in terms of key metrics such as BER, SNR and data rate, as given in Table 1. Table 1 shows some key findings from comparing different modulation schemes in various settings.

Table 1 Comparison of different modulation schemes

Modulation scheme	BER	SNR	Data rate
PWM	Large	Small	Smallest
PPM	Large	Medium	Small
CSK	Smaller	Medium	Medium
OOK	Smallest	Large	Medium
MIMO OFDM	Large	Small	Largest
MPPM	Smaller	Large	Large
VPAPM	Smaller	Large	Large
DIPPM	Smaller	Medium	Large

It is also reported that the power requirement of MPPM is less than that of VOOK and VPPM. The power required for VOOK and VPPM is almost 8 dB more than MPPM to achieve a minimum diming factor of 0.02. The power is also found to increase if the diming factor (D-factor) value is raised from its average value 0.5. It is also noticed that the spectral efficiency of VOOK is better than VPPM with both having the same power. The spectral efficiency of VPPM, VOOK and MPPM is found as 0.02, 0.1 and 0.3, respectively, for the min. D-factor value of 0.02. As a result, MPPM is recommended for superior spectral efficiency results. A D-factor with an inverse relationship to its average value determines the spectral efficiency. VPAPM has a lower power need than overlapping pulse position modulation (OPPM), although it is higher than VPPM and RZ-OOK. VPAPM requires approximately 5 dB of power to achieve the dimming level of 0.5 at the indicated D-factor value $M = 4$. To reach the same dimming level of 0.5, the D-factor values VPPM and RZ-OOK require around 2 dB and 1.5 dB, respectively. It is further demonstrated that VPAPM outperforms VOOK, VPPM, OPPM and RZ-OOK in terms of bandwidth efficiency for the given D-factor values. The VPAPM offers maximum bandwidth efficiency with $M = 4$ at the D-factor value of 0.5. It was also demonstrated that DIPPM needs less power than VPPM with the same BER for all D-factor values. DIPPM gives lowest power at the dimming factor 0.5, which seems to increase when D-factor value increases.

4 Conclusion

The performance of several modulation techniques on VLC is examined in this research. Controlling the dimming level of VLC sources is the most essential component in VLC. As a result, it is preferable to select a modulation scheme that minimizes the effects of dimming while also providing superior data rate results. In terms of power consumption, spectrum efficiency and data rates, we discovered that MPPM

delivers dependable communication comparison with other VLC schemes such as VOOK, VPPM, VPAPM, PWM, DIPPM, NRZ and OFDM.

References

1. Kahn JM, Barry JR (1997) Wireless infrared communications. Proc IEEE 85:265–298
2. Pradana A, Ahmadi N, Adionos T (2015) Design and implementation of visible light communication system using pulse width modulation. In: The 5th international conference on electrical engineering and informatics, 10–11 Aug 2015, Bali, Indonesia
3. Zukauskas MS, Gaska R (2002) Introduction to solid state lighting. Wiley, Hoboken, NJ, USA
4. Rea MS (2000) The IESNA lighting handbook: reference & application. Illuminating Engineering Society North America, New York, NY, USA
5. Berman SM, Greenhouse DS, Bailey IL, Clear RD, Raasch TW (1991) Human electroretinogram responses to video displays, fluorescent lighting, and other high- frequency sources. Optom Vis Sci: official publication of the American Academy of Optometry 68(8):645–662. https://doi.org/10.1097/00006324-199108000-00012
6. IEEE standard for local and metropolitan area networks part 15.7: short-range wireless optical communication using visible light, IEEE Std. 802.15.7, Sep 2011
7. Grubor J, Lee SC, Langer KD, Koonen T, Walewski JW (2007) Wireless high-speed data transmission with phosphorescent white-light LEDs. In: Proceedings of 33rd European conference exhibition optical communication—post-deadline papers (published 2008), Sep 2007, pp 1–2
8. Vucic J et al (2009) 125 Mbit/s over 5 m wireless distance by use of OOK-modulated phosphorescent white LEDs. In Proceedings of 35th ECOC, Sep 2009, pp 1–2
9. Fujimoto N, Mochizuki H (2013) 477 Mbit/s visible light transmission based on OOK-NRZ modulation using a single commercially available visible LED and a practical LED driver with a pre-emphasis circuit. Presented at the optical fiber communication Conference/National Fiber Optic Engineers Conference, Anaheim, CA, USA, 17–21 Mar 2013, Paper JTh2A-73
10. Muthu S, Gaines J (2003) Red, green and blue LED-based white light source: Implementation challenges and control design. In: IEEE 38th IAS Annual Meeting on Conference Record, Oct 2003, vol 1, pp 515–522
11. Ohtsuki T, Sasase I, Mori S (1993) Overlapping multi-pulse pulse position modulation in optical direct detection channel. In: Proceedings of ICC'93-IEEE international conference on communications, vol 2. May 1993, pp 1123–1127
12. Kaur S, Liu W, Castor D, VLC dimming support IEEE 802.15-09-0641-00-0007. [Online]. Available: https://mentor.ieee.org/802.15/dcn/09/15-09-0641-00-0007-vlc-dimming-proposal.pdf
13. Kwon JK (2010) Inverse source coding for dimming in visible light communications using NRZ-OOK on reliable links. IEEE Photon Technol Lett 22(19):1455–1457
14. Georghiades C (1994) Modulation and coding for throughput efficient optical systems. IEEE Trans Inf Theory 40(5):1313–1326
15. Ohtsuki T (2000) Rate-adaptive indoor infrared wireless communication systems using repeated and punctured convolutional codes. 4(2):56–58
16. Bai B, Xu Z, Fan Y (2010) Joint LED dimming and high capacity visible light communication by overlapping PPM. In: Proceedings of 19th annual wireless and optical communications conference WOCC, May 2010, pp 1–5
17. Sugiyama H, Nosu K (1989) MPPM: a method for improving the band-utilization efficiency in optical PPM. J Lightwave Technol 7(3):465–472
18. Ohtsuki T, Sasase I, Mori S (1994) Performance analysis of overlapping multi-pulse pulse position modulation (OMPPM) in noisy photon counting channel. In: Proceedings of IEEE international symposium on information theory, Jun 1994, p 80

19. Ohtsuki T, Sasase I (1994) Error performance of overlapping multipulse pulse position modulation (OMPPM) and trellis coded OMPPM in the optical direct-detection channel. IEICE Trans Commun 77(9):1133–1143
20. Zwillinger D (1988) Differential PPM has a higher throughput than PPM for the band-limited and average-power-limited optical channel. IEEE Trans Inf Theory 34(5):1269–1273
21. T. Ohtsuki, I. Sasase, and S. Mori (1994) Differential overlapping pulse position modulation in optical direct-detection channel. In: IEEE SUPERCOMM/ICC, Conference Record, vol 2. May 1994, pp 680–684
22. Noshad M, Brandt-Pearce M (2012) Expurgated PPM using symmetric balanced incomplete block designs. IEEE Commun Lett 16(7):968–971
23. Afgani MZ, Haas H, Elgala H, Knipp D (2006) Visible light communication using OFDM. In: Proceedings of 2nd International Conference TRIDENTCOM, pp 129–134
24. Armstrong J, Lowery A (2006) Power efficient optical OFDM. Electron Lett 42(6):370–372
25. Elgala H, Mesleh R, Haas H, Pricope B (2007) OFDM visible light wireless communication based on white LEDs. In: Proceedings of IEEE 65th vehicular technology conference-VTC-Spring, Apr 2007, pp 2185–2189
26. Burchardt H, Serafimovski N, Tsonev D, Videv S, Haas H (2014) VLC: beyond point-to-point communication. IEEE Commun Mag 52(7):98–105
27. Elgala H, Mesleh R, Haas H (2009) Practical considerations for indoor wireless optical system implementation using OFDM. In Proceedings of 10th Conference on Telecommunications, Jun 2009, pp 25–29

Simultaneous Estimation of Nebivolol and Cilnidipine in Pharmaceutical Formulation by Reverse-Phase High-Performance Liquid Chromatography Method

Pinal J. Patel, Drashti Mahendrabhai Patel, Meghal J. Patel, Manisha Chaudhari, Kinjal Gandhi, Shashi V. Ranga, and Hardik Mahendrabhai Patel

Abstract In this paper, we have suggested a method for simultaneous quantification using high-performance liquid chromatography for Nebivolol and Cilnidipine in pharmaceutical dosage forms that is accurate, easy, and fast. The separation is achieved using a mobile phase of 50:30:20 v/v acetonitrile, methanol, and potassium dihydrogen orthophosphate buffer, with pH 4.0 adjusted with orthophosphoric acid (10%) on a Phenomenex-luna C18 (250 mm * 4.6 mm, 5) column. We have used a flow rate of e 1.2 mL min^{-1}, with UV detection at 283 nm. Nebivolol and Cilnidipine have retention times of 2.37 and 7.69 min, respectively. For both Nebivolol and Cilnidipine, a linear response is seen over the concentration ranges of 2–10 g/mL ($R_2 =$ 0.998) and 4–20 g/mL ($R_2 = 0.997$). For Nebivolol, the limit of quantitation (LOQ) and limit of detection (LOD) are 0.13 and 0.40 g/mL, respectively, and for Cilnidipine, they are 0.11 and 0.35 g/mL. The percent recovery for Nebivolol is 101.4–101.7%, while for Cilnidipine it is 100.5–100.9%. It is found that the procedure is accurate, exact, linear, sensitive, and robust.

Keywords Nebivolol (NBV) · Cilnidipine (CIL) · Liquid chromatography · Combined dosage form · Method validation

P. J. Patel (✉) · S. V. Ranga
Government Engineering College, Gandhinagar, India
e-mail: pinalpatel@gecg28.ac.in

D. M. Patel
Medusind Solutions, Ramanathapuram, India

M. J. Patel
IBM, Ahmedabad, India

M. Chaudhari
Department of Computer Engineering, Indus University, Ahmedabad, India

K. Gandhi
Parul University, Waghodia, India

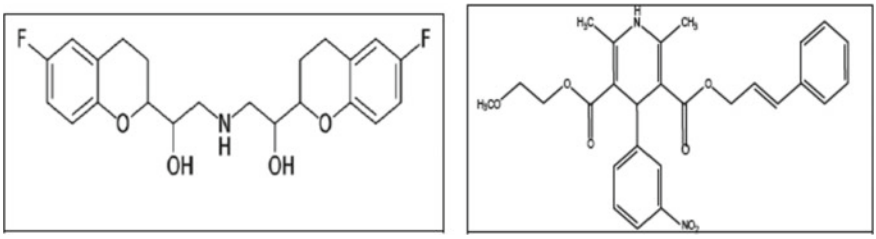
H. M. Patel
Vidush Somany Institute of Technology and Research, Kadi, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_29

1 Introduction

For the treatment of hypertension, a tablet dosage form of a combined fixed-dose formulation including Nebivolol and Cilnidipine is offered. Nebivolol (NBV), also known as '- [iminobis (methylene)] bis [6-fluoro-3, 4-dihydro-2H-1-benzpyran-2-methanol], is an antihypertensive drug that works by inhibiting vascular endothelial nitric oxide release and acting as a 1-antagonist [1]. Nebivolol is official in IP [2]. Cilnidipine (CIL) 1, 4-Dihydro-2, 6-dimethyl-4(3nitrophenyl) 3, 5pyridinedicarboxylic acid 2-methoxyethyl (2E)-3-phenyl-2-propenylester. Cilnidipine (CIL) is a dihydropyridine calcium-channel blocker. It causes vasodilation by inhibiting calcium input into cells. It has greater selectivity for vascular smooth muscle. Cilnidipine is not listed in IP, BP, and USP [3, 4]. Pharmacological rationale is a combined fixed-dose formulation LN βeta5 containing Nebivolol and Cilnidipine. It is a hypertension medication that comes in tablet form. Cilnidipine alone requires 20 mg once daily, and Nebivolol alone requires a maximum dose of 40 mg/day. The combinations of both drugs reduce the dose of the individual drug. A pharmaceutical composition of Nebivolol 5 mg and Cilnidipine 10 mg provides an excellent synergistic effect on lowering blood pressure.

According to the literature, numerous methods (UV, HPLC, and HPTLC) [5–13] have been described for the analysis of individual drugs and combinations of drugs, but there is no method for measuring both Nebivolol and Cilnidipine at the same time that has been documented. As a result, the purpose of this research is to create and validate an RP-HLC method for estimating Nebivolol and Cilnidipine in a combined dosage form (Fig. 1).



(a) Nebivolol (b) Cilnidipine

Fig. 1 Chemical structure of Nebivolol **a** and Cilnidipine **b**

2 Materials and Method

2.1 Reagents and Chemicals

Cadila Pharmaceuticals, Ahmedabad, India, and Nissan pharmaceuticals, Ankleshwar, India graciously have sent pure Nebivolol (NBV) and Cilnidipine (CIL) as a gift sample. Tablet formulation of LN eta 5 (Nebivolol 5 mg and Cilnidipine 10 mg) is obtained from the local market. All of the solvents and reagents are HPLC grade. Merck Pvt. ltd. has provided HPLC quality water, acetonitrile, methanol, and orthophosphoric acid.

2.2 Instrumentation

LC-20AD Prominence Liquid Chromatography with UV detector is used for the chromatographic separation. The output single at wavelength 283 nm is monitored and integrated using Spinchrom CFR software. Weighing is done with an electronic analytical balance (Wensar DA 220). For technique development, a Phenomenex-luna C18 (250 mm * 4.6 mm, 5) column is employed to achieve drug separation at ambient temperature.

2.3 Preparation of Diluent and Mobile Phase

- **Diluent**: Accurately measured 50 volumes of acetonitrile, 30 ml methanol, and 20 volumes of water are mixed and degassed using 0.45 μ membrane filtration assembly.
- **Mobile phase**: Accurately weighed 1.379 g of potassium dihydrogen phosphate is transferred into a 1000 ml beaker containing water, dissolve, and dilute up to the mark with the same. It will give a 10 mM solution of KH_2PO_4 buffer. The mobile phase has contained ACN, methanol, and a potassium dihydrogen orthophosphate buffer that has been adjusted to pH 4 with 10% orthophosphoric acid. For usage, these are filtered through a 0.45 membrane filter and sonicated to remove any gas. The mobile phase is made up of ACN, methanol, and potassium dihydrogen orthophosphate buffer in a 50:30:20 v/v ratio.

2.4 Preparation of Standard Solutions

- **Nebivolol standard stock solution (100 μg/mL)**: Nebivolol (10 mg) is accurately weighed and transported to a 100 mL volumetric flask, where it is dissolved in the

mobile phase (ACN: methanol: potassium dihydrogen orthophosphate buffer in a 50:30:20 (pH 4) ratio) and made up to the mark with the same to yield a standard stock solution (100 g/mL).

2.5 Preparation of Standard Stock Solution of Binary Mixture of Nebivolol and Cilnidipine

- 5 mg of Nebivolol and 10 mg of Cilnidipine are accurately weighed and transferred to a 100 ml volumetric flask, where they are diluted with mobile phase to yield 50 g/ml of Nebivolol and 100 g/ml of Cilnidipine.

2.6 Preparation Sample Solution of Binary Mixture of Nebivolol and Cilnidipine

- A standard solution of 1.0 ml of Nebivolol and 2.0 ml of Cilnidipine is combined, and the volume is adjusted up to 10 ml with the mobile phase to obtain 10 g/ml of Nebivolol and 20 g/ml of Cilnidipine.

2.7 Preparation of Calibration Curve

- The Nebivolol and Cilnidipine calibration curve is made up of various concentrations of binary mixture solution of Nebivolol (50 g/ml) and Cilnidipine (100 g/ml).
- The solution is made by pipetting off 0.2, 0.4, 0.6, 0.8, and 1.0 ml of a binary mixed solution of Nebivolol (50 g/ml) and Cilnidipine (100 g/ml) in a 10 ml volumetric flask and adjusting the volume with mobile phase up to the mark, giving 2–10 g/ml of Nebivolol. The chromatogram is recorded with the chromatographic settings that have been finalized.
- The peak area versus concentration (g/ml) calibration curves for Nebivolol and Cilnidipine are plotted, respectively.

3 RP-HPLC Method for Simultaneous Estimation of Nebivolol and Cilnidipine in Pharmaceutical Formulation

The average weight of 20 tablets is computed, and they are then pulverized in a mortar. A precisely weighed amount of powder equivalent to 10 mg of Cilnidipine or

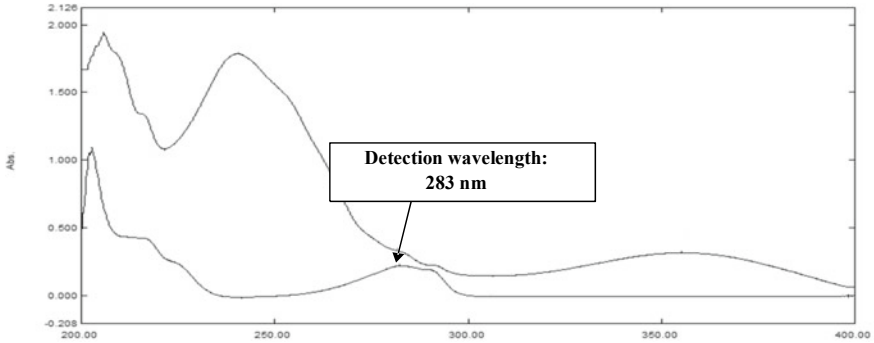


Fig. 2 Overlay UV spectra of Nebivolol (2 μg/ml) and Cilnidipine (17 μg/ml) in methanol showing selection of wavelength detection

5 mg of Nebivolol is transferred to a 100 ml volumetric flask, dissolved in the mobile phase, and finally diluted in the mobile phase, and sonicated for 15 min. Watmann filter paper is used to filter the content. This solution has 50 g/ml NBV and 100 g/ml CIL in it. 1 ml of the aforementioned solution is placed in a 10 ml volumetric flask and diluted to the desired concentration using ACN: methanol: potassium dihydrogen orthophosphate buffer (50:30:20 v/v) pH 4.0 to yield 5 g/ml NBV and 10 g/ml CIL. The resultant solution is filtered using a Millipore filter before being injected into a test tube for analysis. Each drug's concentration is determined using a regression line equation.

3.1 Selection of Wavelength

In methanol, a solution of Nebivolol and Cilnidipine is produced and scanned in the 200–400 nm range. The overlay spectra of Nebivolol (2 μg/ml) and Cilnidipine (17 μg/ml) are used to determine the wavelength. At 283 nm, both Nebivolol and Cilnidipine have a good reaction (Fig. 2).

4 Result and Discussion

4.1 Optimization of Chromatographic Conditions

Acetonitrile: water, methanol: water, phosphate buffer: acetonitrile, and phosphate buffer: methanol are all taken in varying quantities. The 50:30:20 (pH 4) v/v) mixture of ACN, methanol, and potassium dihydrogen orthophosphate buffer gave optimal polarity for efficient migration, separation, and resolution of NBV and CIL. The

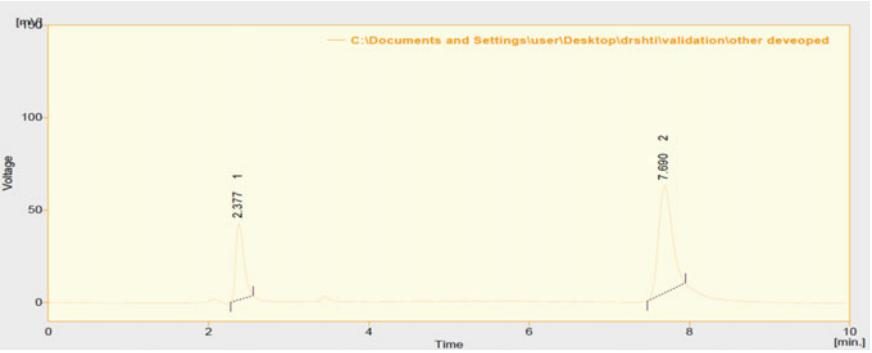


Fig. 3 Optimized chromatogram of Cilnidipine (20 µg/ml) (RT) 7.69 min and Nebivolol (10 µg/ml) (RT) 2.37 min

eluted peaks are clearly defined, resolved, and tail-free under these circumstances. The elution order is NBV ($R_t = 2.37$ min) and CIL ($R_t = 7.69$ min) at a flow rate of 1.2 ml/min (Fig. 3).

4.2 Method Validation [14]

The linearity, range, accuracy, precision, LOQ, LOD, and robustness of the method are validated.

4.3 System Suitability Test Parameter

See Table 1.

Table 1 System suitability test parameter of Nebivolol and Cilnidipine

Parameters	Nebivolol	Cilnidipine	Standard values
Retention time (RT)	2.37 min	7.69 min	–
Tailing factor	1.9	1.5	Not greater than 2.0
Theoretical plates	2299	8765	Greater than 2000
Resolution	–	20.2	Greater than 2

4.4 *Specificity*

- Specificity is a procedure for detecting quantitatively the presence of a component in the sample matrix that is expected to be present. To check for excipient interference, commonly used excipients are spiked in a pre-weighed quantity of medication and the peak area is measured. The absence of excipient interference with the peak of Nebivolol and Cilnidipine is demonstrated by comparing the chromatograms of the blank (mobile phase), standard solution, and test preparation solution. Individual Nebivolol (10 g/ml, R_t = 2.37 min) and Cilnidipine (20 g/ml, R_t = 7.69 min.) chromatograms are then recorded (See Figs. 4, 5, and 6).

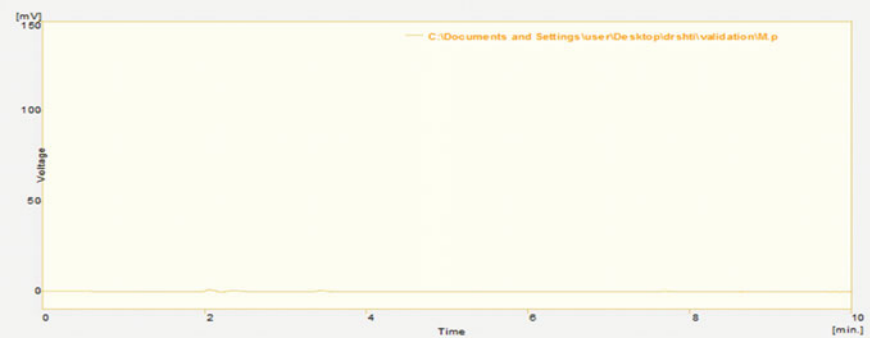


Fig. 4 Specificity chromatogram of blank

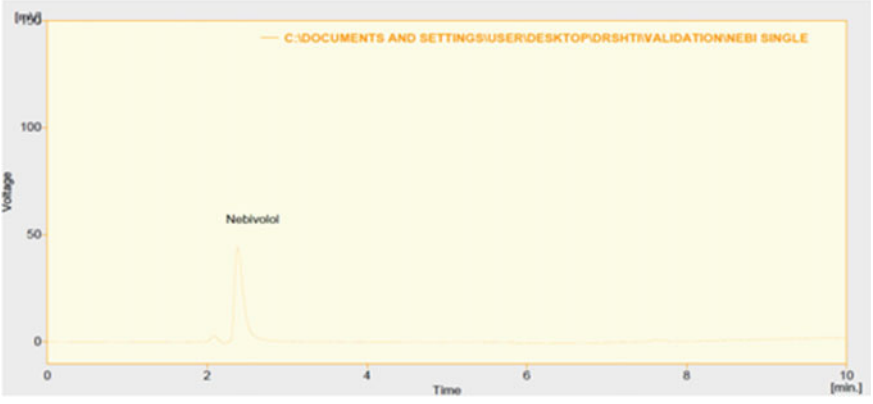


Fig. 5 Specificity chromatogram of standard Nebivolol (10 µg/ml) RT (2.37)

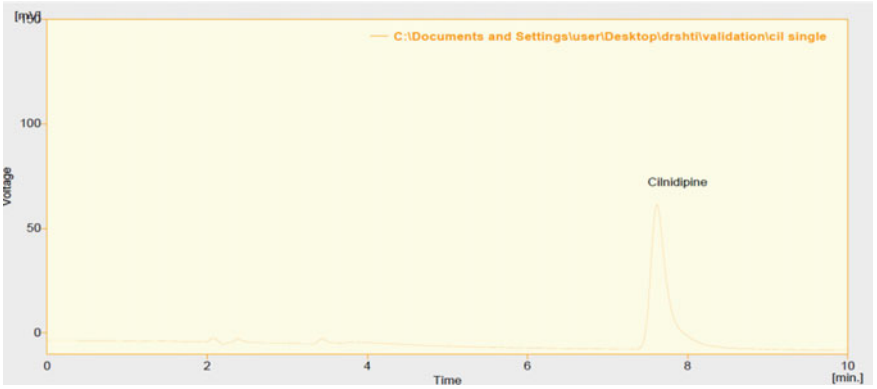


Fig. 6 Specificity chromatogram of standard Cilnidipine (20 µg/ml) RT (7.69)

4.5 Linearity

Both medicines are studied for linearity at five distinct concentration levels. The linearity of Nebivolol and Cilnidipine is 2–10 g/ml and 4–20 g/ml, respectively, as indicated in Tables 2 and 3, and calibration curves in Figs. 7 and 8.

Table 2 Linearity data of Nebivolol

Concentration (µg/ml)	Mean area (mV s) ± S.D. ($n = 5$)	% RSD
2	123.6 ± 1.14	0.92
4	162.8 ± 1.92	1.18
6	212.8 ± 1.92	0.90
8	248.6 ± 2.07	0.83
10	287.2 ± 2.58	0.90

Table 3 Linearity data of Cilnidipine

Concentration (µg/ml)	Mean area (mV s) ± S.D. ($n = 5$)	% RSD
4	273.8 ± 1.92	0.70
8	383.6 ± 2.07	0.50
12	485.2 ± 2.58	0.53
16	595.2 ± 2.86	0.48
20	676.2 ± 2.38	0.35

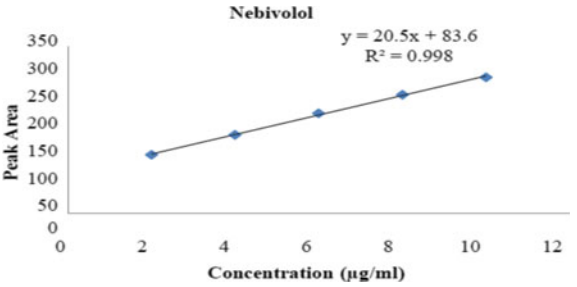


Fig. 7 Calibration curve of Nebivolol (2–10 µg/ml)

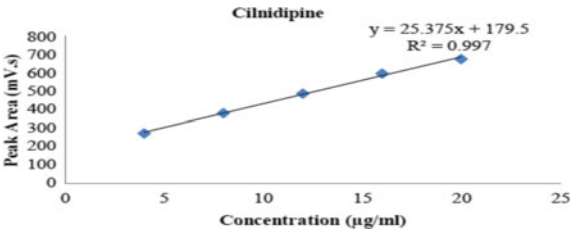


Fig. 8 Calibration curve of Cilnidipine (4–20 µg/ml)

4.6 Accuracy

The method's accuracy was proven by a three-level recovery study from pharmaceutical formulation (80, 100, and 120%). The percentage recovery for Nebivolol was found to be between 101.4 and 101.7%, while Cilnidipine was found to be between 100.5 and 100.9%. Table 4 shows the results, and the percentage recovery was found to be within range.

Table 4 Accuracy data for Nebivolol and Cilnidipine

Drug	Conc. of std drug	Level of recovery (%)	Amount of drug used (µg/ml)	Total amount of the drug (µg/ml)	Amount of drug recovered (µg/ml)	% Recovery ± SD ($n = 3$)	%RSD
Nebivolol	4	80	3.2	7.2	7.30	101.4% ± 0.42	0.41
		100	4	8	8.12	101.5% ± 0.25	0.24
		120	4.8	8.8	8.95	101.7% ± 0.17	0.16
Cilnidipine	8	80	8	14.4	14.48	100.5% ± 0.24	0.24
		100	10	16	16.13	100.8% ± 0.31	0.30
		120	12	17.6	17.77	100.9% ± 0.46	0.45

5 Conclusion

The suggested RP-HPLC method is simple, dependable, and selective, delivering acceptable accuracy and precision with a lower detection and quantification limit. The overall chromatographic run time was 10 min, with eluting times of 2.37 min and 7.69 min for Nebivolol and Cilnidipine, respectively. In pharmaceutical formulation, the assay result was comparable to the labeled amount of each medication. These findings show that the established RP-HPLC technology is easy to use, precise, accurate, and reliable.

References

1. Drug profile of Nebivolol HCl. http://www.drugbank.ca/drugs/db04861. Accessed on 25/11/2014
2. Indian Pharmacopoeia—2014, Government of India Ministry of health and family welfare. Published by Indian Pharmacopoeia Commission, vol 3, pp 2311–2312
3. Drug profile of Cilnidipine. www.mims.com/druginfo/Cilnidipine. Accessed on 26/11/2014
4. The Merck index, an encyclopedia of chemicals, drugs and biological; Cilnidipine-2216; 14th ed. Merck Research Laboratory, 2006, p 379
5. Minase AS, Dole MN, Sawant SD (2014) Development and validation of analytical method for simultaneous estimation of Cilnidipine and olmesartan medoxomil in bulk and tablet dosage form by RP-HPLC. Int J Pharm Pharm Sci 6:508–511
6. Sidhdhapara M, Patel B, Parmar A, Vekariya A, Patel P (2014) Derivative spectrophotometric method for simultaneous determination of Cilnidipine and olmesartan minoxidil in tablet dosage form. Scholars Res Libr 6(3):175–178
7. Mohammed MS (2013) Spectrophotometric method for the estimation of Cilnidipine in bulk and pharmaceutical dosage form. Orient J Chem 29(1):131–134
8. Siddiqui MI, Srinivas M (2014). Simultaneous estimation of telmisartan and Cilnidipine in bulk and tablet formulation using RP-HPLC. Int J Adv Pharm Sci 5:2142–2148
9. Minase AS, Dole MN (2014) Development and validation of analytical method for simultaneous estimation of Cilnidipine and olmesartan medoxomil in bulk and tablet dosage form by HPTLC. J Adv Sci Res 5(3):34–38
10. Sharma T, Rajesh T, Dannana SG, Sudam C (2012) Development and validation of U.V. spectrophotometric method for determination of Nebivolol hydrochloride following ICH guidelines and study of its degradation profile. Asian J Pharm Clin Res 69–72
11. Patil PB, Chavan CB, Jagtap DA, Mohite SK, Magdum CS (2012) Simultaneous estimation of Nebivolol hydrochloride and amlodipine besylate by U.V. spectrophotometric method. Int J Chem Tech Res 4:1241–1246
12. Patel RK, Patel JB (2011) Simultaneous equation spectrophotometric method for the estimation of Nebivolol HCl and hydrochlorothiazide in their combined dosage form. Int J Pharm Appl Sci 2(1):6–10
13. Patel LJ, Suhagia BN, Shah PB (2007) RP-HPLC and HPTLC methods for the estimation of Nebivolol hydrochloride in tablet dosage form. Indian J Pharm Sci 69:594–596
14. The international conference on harmonization of technical requirements for registration of pharmaceuticals for human use, validation of analytical procedure: text and methodology, ICH Q2 (R1), 2005

Security Techniques Implementation on Big Data Using Steganography and Cryptography

Divya Sharma and Ruchi Kawatra

Abstract The COVID-19 pandemic has increased everyone's exposure to the Internet thus there has been an addition of new people who now access the Internet and its applications. Hence, the quantity of big data has grown tremendously in the era of smart city life. Thus concern for security of big data has risen. The attacks currently faced by big databases such as misuse, misrepresentation, modification, and unauthorized users and such attacks have increased in number. To ensure the security of big data stores to prevent misuse, misrepresentation, modification, and unauthorized users issues related to insecurity in storage units and not rendering it vulnerable to attackers, it is needed to enhance the secrecy, privacy, and increase the capacity for hiding secret cover. While cryptography guarantees authentication, integrity, non-repudiation, etc. In this paper, the author has studied popular data hiding techniques, especially steganography and cryptography used for provides security to big databases.

Keywords Vulnerable · Authentication · Smart city · Robustness · Authenticity

1 Introduction

Firstly, let's understand in detail the digital data. Data is the raw information while information is the processed form of data. Digital or electronic data is that data that exists on our computer or the Internet. The various kinds of digital data are: (a) **Digital Audio**: Electronic audio files are stored, transferred, modified, and deleted on your computer, Internet or CD, DVD, etc., such files generally have format .avi, .mp3, etc. (b) **Digital Video**: Digital videos are made up of frames of image synchronized with digital audio digital video files which can also be performed with read, write, update, and modify functions commonly implemented formats include .mp4, .rm, etc. (c) **Digital Text**: Digital text file is made up of alphabets, and these formats are

D. Sharma (✉) · R. Kawatra
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: divya009sharma@gmail.com; divya1007cse.phd21@chitkara.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_30

.pdf, .doc, etc., (d) **Digital Image**: Images are made up of pixels depending upon property it could be gray-white, RGB, and formats are .jpeg, .png, etc.

Let's now understand the commonly used terms in this paper: steganography, cryptography, and big data:

1.1 Steganography

Steganography signal consists of two things: cover signal and secret signal. The art of hiding the secret signal inside a cover signal such that the cover signal seems unchanged is called steganography and the signal generated during steganography is called stego signal. Steganography is currently carried out on the following digital file formats: (a) **Image**, (b) **Audio,** (c) **Video**, and (d) **Text Data**. The commonly popular steganography algorithms are:

- **Spatial domain**: Here pixel of an image is used for hiding secret messages.
- **Frequency domain**: Hiding secret messages in frequency domain of cover signal.

1.2 Big Data

Big data is being created with the creation of new social networking devices, weather forecasting, profile creation (patient, shopping mall, customer profile, etc.) such Web sites create and use big data whether to search, share, update, etc. Big data are often divided into sensitive data and insensitive data [29]. **Sensitive data** such as user name and password are used for authentication of the user while **insensitive data** are those which are common data that do not need to be sensitized.

The sensitive data present here are passwords, social security numbers, credit card details, other personal information, person's identity information, future agendas, political agendas, and confidential corporate information. The areas of application of big data are manufacturing, logistics, transportation, healthcare, government, social networking, weather forecasting, market, finance, smart city data, and its utilities. The big data in health care include X-ray scans, CT scan, MRI scan, patients' history, profile details of patients, also doctor consultation records, etc. (Fig. 1).

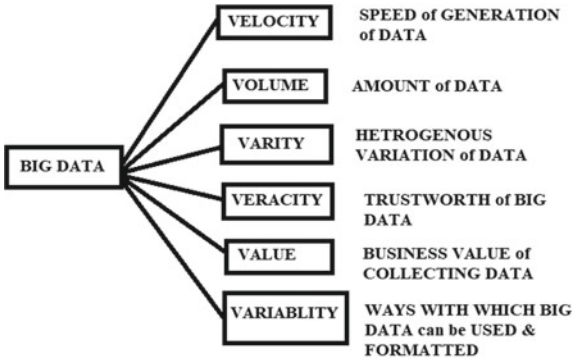
Big data should support properties of easy and cost-effective retrieval, searching, modification, and updating data into it [5]. Big data is divided into three basic types:

- **Structured**: This data exists in a structured form, therefore, in the form of rows or columns or tabular form data created in Oracle, MySQL, SQL, etc.
- **Unstructured**: These data are usually unstructured such as images, audio, etc.
- **Semi-structure**: These include a combination of both structured and semi-structured data. Big data: is a collection of large heterogeneous datasets.

SENSTIVE DATA	NON SENSTIVE DATA
NAME, SURNAME, HOME ADDRESS	NAME OF PLACE, COUNTRY
E-MAIL ADDRESS, LOCATION ADDRES	GENDER, SALALRY DATA OF EMPOLYEE
INTERNET PROTOCOL (IP) ADDRESS	COMPANY WEBSITE, BASIC INFORMATIO
CREDIT CARD DETAILS, ETC	BLOG RECORDS, ETC
PASSPORT, DRIVING LICIENCE, ETC	FITNESS DATA
INSURANCE DETAILS, ETC	WEATHER FORCASTING DATA

Fig. 1 Example of sensitive and non-sensitive text data

Fig. 2 6 V's of big data



- **Big data management system**: These systems are responsible for collecting, integrating, and storing big datasets.
- **Properties of big data**: See in Fig. 2.

1.3 Cryptography

Cryptography is a combination of two processes that are encryption and decryption. Encryption is the process of converting plain signal into a cipher signal while decryption is the process of converting cipher signal into plain signal. Generally, cryptography is of the following types: (1) **Public Key Cryptography**: Here the sender uses the public key of the receiver to encrypt the plain signal while the receiver uses his private key to decrypt the cipher signal. (2) **Private Key Cryptography**: The sender encrypts the plain signal using its private key while the receiver decrypts the cipher signal with the help of the sender's public key, (3) **Hash based**: A common hash table is created which is used to encrypt and decrypt the signal.

- **Properties of cryptography**: See in Table 1.
- **Area of application:** See in Fig. 3.
- **Types of cryptography:**
  These are of two types: (a) symmetric and (b) asymmetric cryptography.

Table 1 Various properties of cryptography

Properties	Description
Authentication	Only authorized user is given access to the message
Integrity	Ensuring that only authorized user can access the message
Non-repudiation	User's cannot refute from the claim of sending the message
Confidentiality	Once the authentic user has sent a message, it has not been modified by the third party
Access control	Only an authentic user can access the message
Availability	An authorized user's should be able to freely access resources without interruption in service

Fig. 3 Cryptography applications

(a) **Symmetric Cryptography**: The same key is used for the encryption and decryption process often called secret key encryption.
(b) **Asymmetric Cryptography**: Different keys are used for the encryption and decryption process. Asymmetric cryptography is classified as follows: (i) Public Key Encryption and (ii) Private Key Encryption.

(i) **Public Key Encryption**: Here the sender encrypts the message with the public key of the receiver while the receiver decrypts the message with their private key.
(ii) **Private Key Encryption**: The sender encrypts the message with their private key and the receiver decrypts the message with the sender's public key.

- **Popular Cryptography Algorithm:** See Fig. 4 and Table 2**.**

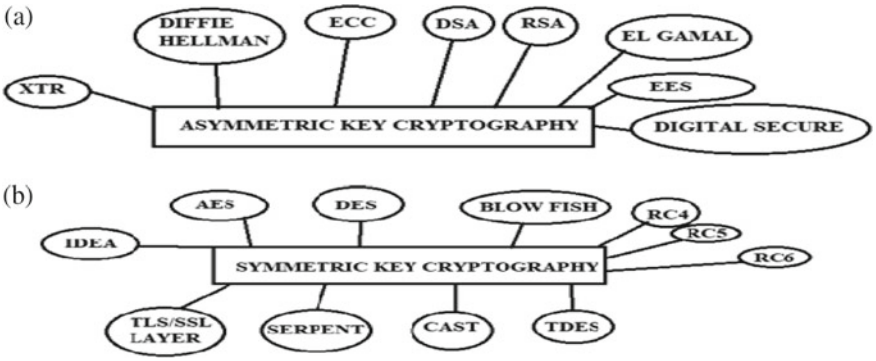


Fig. 4 **a** Different symmetric key cryptography algorithm and **b** different asymmetric key cryptography algorithm

Table 2 Differentiation between symmetric and asymmetric cryptography

Properties	Symmetric cryptography	Asymmetric cryptography
Key	One	Two
Speed of encryption and decryption	Very fast	Slow
Security	Confidentiality	Confidentiality, authentication, and non-repudiation
Resources utilization	Low	High
Example	AES, DES, etc	ECC, DSA, etc
Implementation	Easy	Complex
Transmission	Bulk data	Small and important data
Technique	Old technique	New technique
Used	Most frequently	Many times
Size of encrypted data	Increases	Remains same
Efficiency	Not as efficient as asymmetric	Very efficient
Understandability	Easy to understand	More complex than symmetric

1.4 Cloud Computing

Cloud computing is implemented for storing big data thus increasing the accessibility of big data. Cloud storage is less costly compared to the in-house databases which involve maintenance costs. The data in the big dataset is increasing at an exponential rate especially during COVID-19. Clouds are generally classified as public, private, and hybrid clouds. Commonly implemented clouds are the hybrid clouds which are a combination of both the public and private clouds. Where the data is stored based on their sensitivity in the hybrid cloud. The sensitive data is stored in the private cloud while the insensitive data is stored in the public cloud.

National Institute Standards and Technology (NIST) defined three service models of cloud; these are service as a service (SaaS), cloud platform as a service (PaaS), and cloud infrastructure as a service (IaaS). Cloud data suffers from architecture-related issues, attack vulnerability, and cryptography algorithms (Table 3).

2 Literature Review

The literature review of the paper read by the author has been discussed in Table 4. This tabulated discusses the work done in the cited papers which includes parameters such as data format (like big data, images, audio, etc.) on which data hiding technique has been applied, the data hiding approach that different researchers have used for hiding data such as cryptography or steganography. While the technique mentions the algorithm used for performing the data hiding, along with the advantages and limitations of implementing the same as stated by their respective researchers in their work. This study gives a briefing on the currently popularly implemented security techniques available in the papers and also highlights the parameters across which their properties have been enhanced further. This literature review highlights the algorithms which have been used till date helping future researchers understand the current state and derive a new understanding of this algorithm for big data security.

With the help of the above study, we were able to outline the major properties of cryptography and steganography which are currently been dealt with to provide data security. Thus, observed properties of steganography have been studied (i) capacity of hidden message, (ii) robustness, (iii) tamper resistance, and (iv) transparency similarly we observe (i) authenticity, (ii) integrity, (iii) privacy, and (iv) reliability. With limitations such as Ren et al. [27] were not able to improve the capacity, Shyla et al. [31] the secret message could be destroyed easily all this helps us understand the fundamental need to develop a new algorithm that can improve over the existing lacks in the security of big data.

3 Tools Used for Obtaining Results

A study on the tools and results of the cited paper is obtained and tabulated in Table 5. Mentioning the tools and technology used by the authors of the cited publication along with the result obtained by them, respectively.

The studied articles have mostly implemented steganography and cryptography on big data in MATLAB. In conclusion, the results focus either on enhancing the hiding capacity of the cover signal, or enhancing the secrecy, and privacy of the stego signal.

Table 3 Various approaches along with their properties

Approach	Properties	Application	Tools	Advantages	Limitation
Steganography	**Tamper resistance**: In case of tampering, the signal will be damaged **Robustness capacity**: Huge amounts of information can be hidden inside a cover Invisibility: Imperceptibility **Undetectable**: Make the secret message undetectable in comparison with the original signal **Transparency** **Accurate extraction** **Resistance** **Secrecy**	**Confidential communication**: for secure communication between two parties **Protection** against data alteration **An access** control system in digital content distribution **IPR's**: protection of intellectual property rights in case of theater works, movies, games, papers, etc **E-money**: monitoring the movement of electronic money **Military**: to securely communicate information with border troops **Police communication**: for secure communication between police posts **Detection of tamper-proofing**: in case of tampering, the stego signal will lose its significance and such tampered signal will lose its meaning **Copyrights and protection**: in movies, hidden signals are created based on the buyers and seller, location, and then hidden into the movie at various intervals	Exiftool, Exiv, Imagemagick Outguess, Pngtools, Stegextract, Stegsolve, Steghide, Stegoveritas, Stego-Toolkit, Image-Hide Gif-Shuffle, Quick-Stego, EZ-Stego, OpenStego, S-Tools, Jphs, Wbstego-4, Mp3stegoz, Our secret Omnihide Pro, Crypture, Rstego Ssuite Picsel, Camouflage	**Difficult to detect, Faster approach** **Different techniques** are applied differently to various digital data types like audio, images, and video, etc **Robustness**: increases the robustness of the data **Transparency**: the cover message does not seem to be hiding anything and seems genuine or unchanged on its own **Secrecy, Privacy, Integrity, Protection, Non-repudiation, Confidentiality**	In case of storing huge amount of data it can be easily detected If used wrongly then can prove dangerous In case the algorithm is known then confidentiality is not guaranteed **Modification attack**: in case the parameters of the image are changed then a secret message hidden in the picture may be lost forever **Image rotation**: when the image is rotated then the hidden message may easily be lost **Overhead**: few algorithms generate a larger overhead to hide small secret messages **Done Differently**: steganography is performed differently on different data

(continued)

Table 3 (continued)

Approach	Properties	Application	Tools	Advantages	Limitation
Big data	**Velocity**: how fast our structure or unstructured data in a big data base can be stored or retrieved **Veracity**: the data saved in big data base is heterogeneous data **Volume**: how fast the vast amount of data stored in a big database can be accessed **Value**: the value of the data being stored in case of common data **Variability**: various datasets can be used by the same user or only one dataset could be used by users of the same Social networking websites **Variety**: it is the type of data that is stored in the big database like audio, video, etc. this could also be classified as structured, semi-structured, unstructured data	**Social networking websites**: like facebook, twitter where billions of users access, share and comment various kinds of data **Weather Forecasting** **Profile** creation (patient, shopping mall, customer profile, etc.) such websites create and use big data whether to search, share, update, etc **Banking and security, Communication media and entertainment, Health provider, Education, Manufacturing and natural resources, Government, Insurance, Retail and wholesale, Transportation, Energy and utility**	Apache hadoop, atlas.ti, MoongoDB, HPCC, Apache storm, Qubole, Cassandra Stats iQ, Couch DB, Pentaho, Flink, Cloudera, Openrefine, Rapidminer, Data cleaner, Kaggle, Hive, Spark, R, Data lakes, NoSQL, Apache Grill, Dryad, Jasper soft, Splunk	**Decreases cost**: as various kinds of data are stored in a single location **Increased storage**: the storage capacity of big data stores is greater than traditional databases **Fast storage**: The speed of storing data in a big database is faster **Retrieval of data**: Fast retrieval of the stored data **Flexible storage**: the storage capacity can be increased depending on the requirement **Better performance**: throughput of data retrieval is very high **Concurrent read and write**: big database data can be read and written to multiple users at the same time	**Limited user can lead to biases**: only the data of those user who have accessed the service will be available User level execution only exists in **selected channels**: as the type of data stored in big databases is vast and only some data is generally used for market analysis, the preferences of users there this data needs to be filtered **Modification** of data is required to get accurate results: since data like an image can be shared among many users thus the results are often needed to be aggregated so as to make to more understandable **Algorithms** implemented at the user level are difficult: to analyze the data stored in a big database is to create smaller datasets and then analyze them or analyze the whole dataset which is done by writing a complex algorithm **Big data is not suitable for learning**: practically implementing big data for small businesses is not suitable and not efficient in giving results **Big data often contain noise**: big datasets contain noise that can lead to misleading results and cleaning them is necessary **Not easily accessible or transferable**: to ensure security the big database should not be easily accessible to just anyone. Big data transfer should take place on a secure channel

(continued)

Table 3 (continued)

Approach	Properties	Application	Tools	Advantages	Limitation
Cryptography	**Authentication**: only the legitimate user can access the data **Integrity**: the encrypted data is the same as the original signal hence not been modified **Data repudiation**: **Confidentiality**: increases the confidence of the receiver on the signal received Access control: Availability: the encrypted data can be made available for all users **Privacy**: ensures that the data received is the same as the data sent **Reliability**	**Crypto-currency**: to monitor the flow of crypto-currency and third-party sources from manipulating and introducing their own on the internet **Military**: to provide safe and secure communication between military troops geographically located at different locations **Time stamping**: to save details and credentials of individuals accessing confidential information **Electronic money**: while performing banking transactions securely **Whatsapp and social networking sites**: to securely communicate using what app or any social networking website or applications **Email**: to ensure authenticity, integrity, non-repudiation of the email when sent across the internet **Data privacy**: where there is a need for ensuring the privacy of data when sending it on the internet **Integrity**: therefore all those applications which guaranteed the integrity of the message when on the internet **Banking**: to stop money theft **Ecommerce**: performed by eCommerce websites to provide security to their customers **Datastores**: to provide data stored on the server or PC **Safe internet browsing**: to ensure a secure browsing experience on the internet **Certificates**: the certificates are shared and saved in encrypted form **Authentication**: to ensure the authenticity of encryption is performed	Docker CertMgr.exe SignTool.exe Security token Key-based authentication SignTool Java cryptography architecture	**Authentication**: cryptography ensures that only authentic users can access the information **Confidentiality**: only legitimate users can access the information **Data integrity** ensures that the data received is the same as data sent by the sender **Non-repudiation**: **Safe internet browsing** **Certificates** **Secrecy** **Privacy**	**Non usable**: in case of a system failure of the legitimate user either the complex key or the complex algorithm is lost and hence decrypting them could increase time **Vulnerable**: in the case of a weak algorithm the encrypted signal could be easily broken and tampered with **Cost**: the cost of writing and coding a cryptography algorithm is very large **No hiding**: cryptography does not hide the message but makes it nonunderstandable **Suspicious**: since the message becomes nonunderstandable therefore seems suspicious to some and is prone to deletion attack

(continued)

Table 3 (continued)

Approach	Properties	Application	Tools	Advantages	Limitation
Cloud computing	**On demand service**: cloud services are demand-driven and do not need human intervention for monitoring **Scalability**: Cloud services are scalable thus the capacity of the cloud can be increased or decreased according to requirement **Resource sharing**: cloud allows users separated geographically to share resources **Resource utilization**: the cost of cloud computing is made by monitoring the usage of the resource monitored by the cloud provider	**Art**: the artist can create their art pieces and share the same with the help of cloud **Business**: business data that should be accessible 24*7 by anyone is kept in the cloud **Back up stores**: various backup of application program interfaces are kept in the cloud **Education**: cloud provides a platform for storing lectures and sharing information among the students, especially in the case of distance education **Entertainment**: cloud provides an environment IN which game players from all around the world can come together and take part in a game and the same in the case of video conferences **Management**: data stored in the cloud can easily be managed with help of inbuilt tools, **Social**: helps various people connect	Microsoft Azure, AWS, ALIBAB Cloud, Google Cloud Cloud Monitoring Tools: CLOUD WATCH, Cloud Monitoring Tools, Infrastructure manager, App dynamics, Relics True Sight Pulse, Solar winds, Retrace, Sematext,Aternity	**Backup**: data that has been stored in the cloud in case of damage can be restored as backup is maintained **Data sharing**: data stored on the cloud can be easily shared and accessed by users from anywhere as it is on the internet **Data accessibility**: data stored in cloud could be accessed anytime and anywhere **Less maintenance cost**: data stored on the cloud has very less hardware and software maintenance costs **Portable**: data on the cloud can be accessed on android, tabloid, PC, etc **The cost** of cloud is as per the user's model: the cloud is an application program interface thus the cost of cloud is as per their usage data **Unbound storage**: huge amounts of data files can be stored in a cloud **Security**: cloud provides secure storage and handling of data	**Always connected to the internet**: to access the data on the cloud the user needs a good internet connection **Mobility in cloud**: when mitigating from one cloud service provider to another the data storage model needs to remain the same else users will not be able to access certain data **Control**: the main control over the data on the cloud is with its service provider **Security**: while mitigating data from in-house storage to the cloud the organization's information could be hacked

Table 4 Literature study of the cited papers for data hiding

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
1	Tan et al. [32]	Cloud data	Steganography	Shamir Secret Sharing Algorithm and STC Chain	Ensure privacy and confidentiality Blockchain enhances the data characteristics of unforgivable, traceable, open and transparent, and collective maintenance Prevent attackers from stealing data, but also protect the privacy of keys	The key is broken into fragments Same limitation as symmetric key
2	Al Hamid et al. [2]	Medical Big Data	Cryptography	Illusion-based technique thus Fog Computing and Bilinear Pairing Function and Bilinear Diffie–Hellman Problems and Elliptic Curve Diffie–Hellman (ECDH)	Ensure security of patient medical data	If the decoy seems fishy then the attacker will be able to break the method Only two options are available for choosing an elliptic curve for pairing-based cryptography

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
3	Atiewi et al. [4]	Data on cloud	Cryptography	Fiestel scheme and RC6 and AES and XOR's	Performs authentication, verification, or both thus used by authorized and verified users Provide high security via a gateway device	Could be enhanced further for mutual authentication between gateway devices and IoT devices Stored data in raw form thus it can be easily forged by malicious users

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
4	Rabanal and Martínez [25]	Big Data environment	Cryptography	Adversarial Neural Cryptography	Robustness against attacks Higher system throughput and Reliability Flexible schemes Improve efficiency	Higher cost of implementing complexity algorithm Risks of information leakage in case of targeted attacks Requirement for more flexible schemes Increasing the complexity thus impacts performance In adversarial neural cryptography, here the attacker tries to only decrypt 50% of the encrypted data, the remaining 50% of the data is decrypted correctly then by switching them accordingly as previously done

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
5	Gupta et al. [8]	Big data	Asymmetric Cryptography	Diffie–Hellman and RSA and Elliptic Curve Cryptography (ECC)	Provides a scalable, reliable, and fault-tolerant model Computations are very simple thus more practical solutions to secure big datasets Less complex implementation and maintenance	Efficiency and security were concluded theoretically by comparing it to its peer algorithms and not practically ECC key size is small thus it would be easy to crack
6	Rajaprakash et al. [25]	Big data	Cryptography	New algorithm Rajaprakash Bagathbasha Jaishankar25 (RBJ25)	Increases security	RBJ25 Encryption and decryption performance in term of speed and memory is more than ChaCha, and AES

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
7	OzCakmak et al. [23]	Big data	Cryptography	Quantum Cryptography (Heisenberg's Uncertainty Principle and Photon Polarization Principle) and Neural Cryptography (Artificial Neural Network)	Random secret key High security Knowing the key is insufficient, the attacker should also know the weight and network architecture	High error rate in high bit Works in limited space With increase in data volume, the security of the cryptography becomes risky As based on behavior analysis thus pose a problem for real world big data No action has been taken to protect the quality of big data
8	Yi et al. [36]	Audio	Steganography	AdaMP3Stego (Adaptive Steganography) and Syndrome Trellis Codes (STCs)	Minimize the content distortion Optimized data concealment Format mp3 is preserved	Length of Huffman code stream quantized coefficients cannot exceed the available bits

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
9	Bracciale et al. [5]	Big data	Cryptography	TOOLATE Cryptosystem and SHA256 and KMS	Periodic re-encrypts entire data store using dynamically generated keys in a hash chain Suitable for non-tamper proof and low-power devices Data confidentiality Expired key will be treated the same as a invalid key	Implemented on offline data store In case, the attacker has a valid key which will be valid in the future then by tampering with the device and easily force re encryption which gives access to data
10	Golovnev et al. [7]	Data structure	Cryptography	Fiat–Naor Inversion and 3SUM-Indexing	Efficient and Randomized cryptography method	Hardness observed in space and time complexity with the change in size of data set
11	Maata et al. [20]	Big data	Symmetric Cryptography Algorithm	Two-fish and Feistel Structure	Increases accuracy and efficiency	Space and time for encryption and decryption process have increased significantly

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
12	Hachaj et al. [9]	Images	Steganography	Eigen Algorithm (Principal Component Analysis (PCA)) and Log Euclidean Distance (LED)	Bit-level accuracy obtained Scalable Format of the image (jpeg) is maintained Very good robustness for upscaling A robust approach especially against compression attack	All images need to be cropped into a specific resolution (70*109) face making the eyes of each person-centered and thus the same position Same as the original Eigenfaces approach Not good against image rotation attack
13	Devi and Chamundeeswari [26]	Health care big data	Cryptography	A3DES algorithm (Anonymization technique and Triple Data Encryption Standard) and SHA-256	Reliability Longer key length Diminishes the quantity of time Flexibe and compatible Authentication Better performance than AES and DES. While the file size file is in MB	Same limitation as the symmetric key algorithm The accuracy of the proposed algorithm can be improved further

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
14	Krishna [15]	Big data	Fully Homomorphic Cryptography	ECC (Cubic Spline Curves)	Asymmetric block cipher mode The key length needed for sufficient security of data is less than the key length for the ECC algorithm Consumes less computing resources for processing ciphertexts The encryption process is free from differential side-channel attacks The amount of data stored is large	Authentication not ensured Encryption of data conducted at character level of plain text
15	Adnan and Ariffin [1]	Big data	Cryptography	3D-AES and Rotation Key Function and S-box method	Sensitive data protects data integrity against confidentiality Implements similar operations to AES Higher level of security due to increased bit size Better complexity, security, and performance compared to AES	Same limitation as symmetric key Failed randomness test Simplistic in nature Easily de-constructed or reverse engineered

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
16	Shyla et al. [31]	Image	Steganography	Genetic Algorithm (Selection and Embedding)	Data authentication Effective data integrity Better performance There is the possibility of retrieving the payload data using this algorithm	Embedded payload data could be damaged by the intruders during transmission Compares it with the predefined image in its database and reconstructs the complete payload image
17	Al-Shaarani and Gutub [3]	Image	Cryptography and steganography	LSB and DWT (Steganography) and XOR cipher (Cryptography)	Better robustness and security Secrecy of the secret data	Can be extended for audio and video Could include secret sharing schemes based on random bit shifting method
18	Lingamallu and Veeramani [16]	Image	Steganography and cryptography	Wavelet Transform and LS (Logical and Statistical)	Tackle various attacks like noise, rotational, histogram, and visual attacks Has the ability to counteract the assaults when compared with previous methods Payload capacity increases	Reduces length of ciphertext

(continued)

Table 4 (continued)

S. No	Author	Data format	Data hiding approach	Technique implemented	Advantage	Limitation
19	Ren et al. [27]	Audio	Steganography	Advanced Audio Coding (AAC) (Huffman coding) and STC (syndrome trellis codes) and reverse mapping	Improve statistical security and characteristic Imperceptibility Improve stegano analysis's ability to 8% − 30% Improve auditory concealment	There is no change in capacity of data that can be hidden
20	Liu et al. [18]	Video	Steganography	Intra-Embedding (DCT) and Pre-Embedding (LSB) and Post-Embedding (AES)	Higher embedding capacity Imperceptibility Simple and has a high embedding capacity	The amount of data that can be hidden in video is limited The amount of data embedded in the video is limited Not robust enough Non reversible steganography
21	Liu et al. [19]	Image	Steganography	Coverless Steganography and DenseNet (Convolution Neural Network (CNN) and DWT	Good robustness Better security performance Visual detection of abnormalities from the human perspective is difficult Anti-Detectability	More time is required for comparison to find the perfect cover Limited image set used Robustness could be improved against geometric attacks

Table 5 Tools and technology used by the cited articles and results obtained

Cited as	Technology/system specification used	Result obtained
Sarosh et al. [28]	MATLAB R2017a, Database OPENi	The image is 256 × 256 pixels Encrypted images are shared using the (2, 2)-CSIS
Lingamallu and Veeramani [16]	Written in 'C' language and implemented in MATLAB R 2019b, USC-SIPI images database	Image Lena, Capacity (Bits) = 9424; PSNR = 63.12, CORR = 1.000; NPCR = 0.0104
Al-Shaarani and Gutub [3]	MATLAB environment using randi() function	1-bit and 2-bit share generation methods with various target key sizes starting at 32 bits and up to 3072 bits MSE values close to zero
Shyla et al. [31]	MATLAB 2017a; using USC-SIPI image database set	Images of 512 × 256 Car proposed method = 79.76, SSIM = 0.9979
Adnan and Ariffin [1]	NIST Statistical Test Suite	When r (round) = 3, plaintext size = 16 * 4 bytes to produce a 64-byte output ciphertext with secret key size is 16-byte
Tan et al. [32]	Blockchain in Alibaba Cloud Ubuntu 18.04, configured with 2-core, 16G, and 100G storage and applied XuperChain-crypto library module	Split time ≈ 150 ms, reconstruction time ≈ 2 ms when $n = 10$ and $t = 6$ Split time ≈ 100 ms, reconstruction time ≈ 1 ms when $n = 7$ and $t = 3$ Split time ≈ 330 ms, reconstruction time ≈ 22 ms when $n = 20$ and $t = 12$

4 Conclusions

In the current era, there is a need for enhancing the secrecy and privacy of big data as well as the capacity of the cover signal for hiding a bigger secret message. The author proposes the use of steganography techniques on big data such that the resulting stego media contains no detectable artifacts. The resultant could be transmitted over the Internet, or saved in the database. When a secret message is needed then it could be extracted from the stego signal with the help of the key. This ensures the secrecy of the message stored inside the data and enhances the trust and authenticity of the end-user. While implementing it on an existing service, the initial time of creating a stego signal will be more thus the time complexity may be more depending on the size of the dataset. The proposed algorithm is designed to improve time complexity and space complexity.

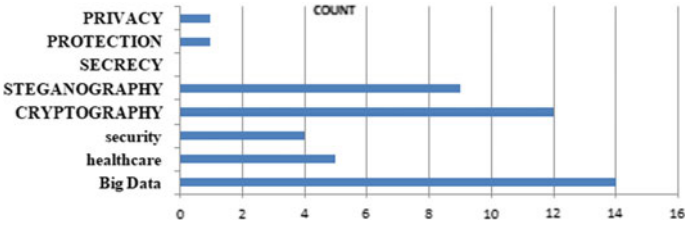


Fig. 5 Number of parameters implemented in the cited papers

Currently, implemented data hiding techniques that provide security to big data have been tabulated in the literature review which helps figure out the various properties where big data needs to be enhanced concerning security. The basic objective found was to hide data to ensure the secrecy and privacy of the data. The main purpose of this paper is to find the objectives that are increasing the payload capacity, making data imperceptible, and increasing the robustness of the data. Therefore, there is a need to improve the capacity of the cover to hide a more secret message and make the cover seem imperceptive without changing the human perception while the robustness is to achieve the hidden message. Also, the privacy and secrecy of the big data when being accessed and while saving from the third party or attacker. Studying the currently popular tests such as mean square error (MSE), peak squared error ratio (PSNR), and correlation analysis (CORR). These tests are used to check the addition of noise to the original signal. Other than these tests, the size of the data after encryption and decryption should be measured along with the time taken for performing encryption and decryption to check for the hardness of the problem. Thus, the data hiding algorithm should make the data imperceptible to the users with unauthorized access. While the data hiding algorithm should be performed in real time thus should not result in a hardness problem (Fig. 5).

References

1. Adnan NAN, Ariffin S (2019) Big data security in the web-based cloud storage system using 3D-AES block cipher cryptography algorithm. In: Yap BW, Mohamed AH, Berry MW (eds) Soft computing in data science, vol 937. Springer Singapore, pp 309–321. https://doi.org/10.1007/978-981-13-3441-2_24
2. Al Hamid HA, Rahman SMM, Hossain MS, Almogren A, Alamri A (2017) A security model for preserving the privacy of medical big data in a healthcare cloud using a fog computing facility with pairing-based cryptography. IEEE Access 5:22313–22328. https://doi.org/10.1109/ACCESS.2017.2757844
3. Al-Shaarani F, Gutub A (2021) Securing matrix counting-based secret-sharing involving crypto steganography. J King Saud Univ Comput Inform Sci S1319157821002603. https://doi.org/10.1016/j.jksuci.2021.09.009
4. Atiewi S, Al-Rahayfeh A, Almiani M, Yussof S, Alfandi O, Abugabah A, Jararweh Y (2020) Scalable and secure big data iot system based on multifactor authentication and

lightweight cryptography. IEEE Access 8:113498–113511. https://doi.org/10.1109/ACCESS.2020.3002815
5. Bracciale L, Loreti P, Raso E, Bianchi G (2021) Toolate: cryptographic data access control for offline devices through efficient key rotation. In: Proceedings of the 2th workshop on CPS&IoT security and privacy, pp 57–62. https://doi.org/10.1145/3462633.3483982
6. Daraghmeh R, Brown R (2021) A big data maturity model for electronic health records in hospitals. In: 2021 international conference on information technology (ICIT), pp 826–833. https://doi.org/10.1109/ICIT52682.2021.9491781
7. Golovnev A, Guo S, Horel T, Park S, Vaikuntanathan V (2020) Data structures meet cryptography: 3SUM with preprocessing. In: Proceedings of the 52nd annual ACM SIGACT symposium on theory of computing, pp 294–307. https://doi.org/10.1145/3357713.3384342
8. Gupta S, Vashisht S, Singh D, Kushwaha P (2019) Enhancing big data security using elliptic curve cryptography. In: 2019 international conference on automation, computational and technology management (ICACTM), pp 348–351. https://doi.org/10.1109/ICACTM.2019.8776764
9. Hachaj T, Koptyra K, Ogiela MR (2021) Eigenfaces-based steganography. Entropy 23(3):273. https://doi.org/10.3390/e23030273
10. Hussain M, Wahab AWA, Idris YIB, Ho ATS, Jung K-H (2018) Image steganography in spatial domain: a survey. Sig Process Image Commun 65:46–66. https://doi.org/10.1016/j.image.2018.03.012
11. Jiang C, Huang C, Huang Q, Shi J (2021) A multi-source big data security system of power monitoring network based on adaptive combined public key algorithms. Symmetry 13(9):1718. https://doi.org/10.3390/sym13091718
12. Kadhim IJ, Premaratne P, Vial PJ, Halloran B (2019) Comprehensive survey of image steganography: techniques, evaluations, and trends in future research. Neurocomputing 335:299–326. https://doi.org/10.1016/j.neucom.2018.06.075
13. Kawatra R, Kumar V, Madan S (2018) Embedding stego text in a digital cover using Elliptic Curve. J Adv Res Dyn Control Syst 11:836–841. ISSN 1943-023X
14. Kawatra R, Kumar V, Madan S (2019) Hiding information along fractal in a digital cover to improve capacity. Published in the Springer book series of Lecture Notes in Electrical Engineering. ICETIT 2019. Scopus indexed. https://doi.org/10.1007/978-3-030-30577-2_93
15. Krishna AVN (2018) A big–data security mechanism based on fully homomorphic encryption using cubic spline curve public key cryptography. J Inf Optim Sci 39(6):1387–1399. https://doi.org/10.1080/02522667.2018.1507762
16. Lingamallu NS, Veeramani V (2021) Secure and covert communication using steganography by Wavelet Transform. Optik 242:167167. https://doi.org/10.1016/j.ijleo.2021.167167
17. Liu H-C, Chen W (2020) Optical ghost cryptography and steganography. Opt Lasers Eng 130(106094):106094. https://doi.org/10.1016/j.optlaseng.2020.106094
18. Liu Y, Liu S, Wang Y, Zhao H, Liu S (2019) Video steganography: a review. Neurocomputing 335:238–250. https://doi.org/10.1016/j.neucom.2018.09.091
19. Liu Q, Xiang X, Qin J, Tan Y, Tan J, Luo Y (2020) Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping. Knowl-Based Syst 192:105375. https://doi.org/10.1016/j.knosys.2019.105375
20. Maata RLR, Cordova RS, Halibas A (2020) Performance analysis of twofish cryptography algorithm in big data. In: Proceedings of the 2020 9th international conference on software and information engineering (ICSIE), pp 56–60. https://doi.org/10.1145/3436829.3436838
21. Muneeshwari P, Athisha G (2020) Extended artificial immune system-based optimized access control for big data on a cloud environment. Int J Commun Syst 33(13):e3947. https://doi.org/10.1002/dac.3947
22. Oussous A, Benjelloun F-Z, Ait Lahcen A, Belfkih S (2018) Big data technologies: a survey. J King Saud Univ Comput Inform Sci 30(4):431–448. https://doi.org/10.1016/j.jksuci.2017.06.001
23. OzCakmak B, OzbIlen A, YavanoGlu U, CIn K (2019) Neural and quantum cryptography in big data: a review. In: 2019 IEEE international conference on big data (big data), pp 2413–2417. https://doi.org/10.1109/BigData47090.2019.9006238

24. Rajaprakash S, Bagath Basha C, Muthuselvan S, Jaisankar N, Singh RP (2020) RBJ25 cryptography algorithm for securing big data. J Phys Conf Ser 1706(1):012146. https://doi.org/10.1088/1742-6596/1706/1/012146
25. Rabanal F, Martínez C (2020) Cryptography for big data environments: current status, challenges, and opportunities. Comput Math Methods 2(1). https://doi.org/10.1002/cmm4.1075
26. Ramya Devi R, Vijaya Chamundeeswari V (2020) Triple DES: privacy preserving in big data healthcare. Int J Parallel Prog 48(3):515–533. https://doi.org/10.1007/s10766-018-0592-8
27. Ren Y, Cai S, Wang L (2021) Secure AAC steganography scheme based on multi-view statistical distortion (SOFMVD). J Inform Sec Appl 59:102863. https://doi.org/10.1016/j.jisa.2021.102863
28. Sarosh P, Parah SA, Bhat GM, Muhammad K (2021) A security management framework for big data in smart healthcare. Big Data Res 25:100225. https://doi.org/10.1016/j.bdr.2021.100225
29. Sharma S, Mehrotra S (2022) Cryptographic techniques for data processing. In: 5th joint international conference on data science and management of data (9th ACM IKDD CODS and 27th COMAD), pp 344–347. https://doi.org/10.1145/3493700.3493771
30. Shi R, Zhang M (2019) Strong privacy-preserving two-party scalar product quantum protocol. Int J Theor Phys 58(12):4249–4257. https://doi.org/10.1007/s10773-019-04296-0
31. Shyla MK, Kumar KBS, Das RK (2021) Image steganography using genetic algorithm for cover image selection and embedding. Soft Comput Lett 3:100021. https://doi.org/10.1016/j.socl.2021.100021
32. Tan L, Yu K, Yang C, Bashir AK (2021) A blockchain-based Shamir's threshold cryptography for data protection in industrial internet of things of smart city. In: Proceedings of the 1st workshop on artificial intelligence and blockchain technologies for smart cities with 6G, pp 13–18. https://doi.org/10.1145/3477084.3484951
33. Vij A, Saini S, Bathla R (2020) Big data in healthcare: technologies, need, advantages, and disadvantages. In: 2020 8th international conference on reliability, infocom technologies and optimization (trends and future directions) (ICRITO), pp 1301–1305. https://doi.org/10.1109/ICRITO48877.2020.9197775
34. Wang Y, Yang K, Yi X, Zhao X, Xu Z (2018) CNN-based steganalysis of Mp3 steganography in the entropy code domain. In: Proceedings of the 6th ACM workshop on information hiding and multimedia security, pp 55–65. https://doi.org/10.1145/3206004.3206011
35. Wang Y, Li Z, Wong DW-C, Cheng C-K, Zhang M (2018) Finite element analysis of biomechanical effects of total ankle arthroplasty on the foot. J Orthop Transl 12:55–65. https://doi.org/10.1016/j.jot.2017.12.003
36. Yi X, Yang K, Zhao X, Wang Y, Yu H (2019) AHCM: adaptive huffman code mapping for audio steganography based on psychoacoustic model. IEEE Trans Inf Forensics Secur 14(8):2217–2231. https://doi.org/10.1109/TIFS.2019.2895200

Identification of Generative Adversarial Network Forms, Open Issues, and Future Study Areas: A Study

Dawit Milkiyas Benti, Shaik Janbhasha, and Eshetu Gusare Desisa

Abstract Generative Adversarial Network is an emerging technology and research area in machine learning from the time 2014. Generative Adversarial Networks (GANs) is a novel class of deep generative models that has recently gained significant attention. GANs learn complex and high-dimensional distributions implicitly over images, audio, and data. Though carry countless exhilarating prospects, these solicitations likewise increase the mindfulness of the hazard of bogus imageries which may basis enormous destruction. This work highlights some key GAN forms, problems, and research gaps in this study. In addition, we address the benefits that GAN could provide to humans as well as potential solutions. To end with, centred on the perceptions enlarged, we extant encouraging study advices in this hastily emergent arena.

Keywords Deep model · Generative Adversarial Networks · Supervised learning · Unsupervised learning

1 Introduction

Preceding towards the advance of profound learning, shadow machine learning approaches, such as SVM and random forests, dominated the computer vision industry. Deep learning models regained dominance with the back propagation technique and the quickening of calculation by Graphics Processing Units, outperforming humans in competitions such as the well-known Image Net. While deep learning has proved successful in a variety of applications, including facial recognition, object detection, traffic prediction, and trade prediction, it has been plagued by the problem that profounder neural nets are more difficult to train and generate minimal progression [1] (Fig. 1).

Generative Adversarial Networks (GAN) is a deep learning system that trains two models at the same time: a generative type G and a discriminative style D. G's

D. M. Benti · S. Janbhasha (✉) · E. G. Desisa
Department of Computer Science, Assosa University, Assosa, Ethiopia
e-mail: afreen.jbasha@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_31

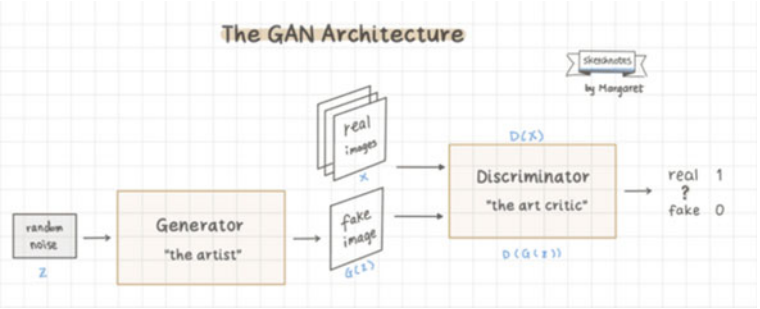


Fig. 1 Architecture of GAN

goal is to figure out how some target data is distributed (e.g., distributions of pixel intensity in images). D facilitates G's training by comparing the data created by G to "actual" data, allowing G to understand the distribution that underlying the real data. Good fellow et al. (2014) defines GAN as a pair of basic neural networks. In practice, though, the models can be any generative-discriminative pair [2].

2 Forms of GANs

2.1 Conditional GAN

If both the Generator and Discriminator are habituated on specific windfall info, GAN can be prolonged to a restricted design. Conditioning can be done by putting y into the Discriminator and Generator as an extra infusion stratum. In their 2014 paper "Conditional Generative Adversarial Networks," Mehdi and Simon originally described the cGAN. The authors of the study justify their method by stating that they want to control the generator model's image production process [3].

2.2 Vanilla GAN

G is a generator that captures the distribution of data (make realistic images). Assessment the chance that an illustration originates commencing the training facts rather than Generator using Discriminator D (tell real and fake images apart). $pz(z)$: noise in the input, pdata(x): the distribution of real data G's generator's distribution, and (pg (x)) (z): minimax game for two players (min G max D). $V(D, G) = Ex\sim$pdata(x) [log $D(x)$] + $Ez\sim pz(z)$ [log(1 − $D(G(z))$)] [4].

2.3 DcGAN

Deep Convolutional Generative Adversarial Neural Network for sensing several innovative arising illnesses in the health arena for which data set is precise restricted, and DCGAN acts as an useful method for creating artificial information because facts for diagnosing Coronavirus is scarce. It also refer to data sets obtained commencing a variety of openly accessible springs, allowing a significant number of picture data sets to be qualified and verified in a variety of newest technology fully convolutional styles to assess each accurateness, intricacy, and processing time [5].

2.4 CatGAN

CatGAN is a category-aware GAN that combines an efficient category-aware style for class text procreation with a priestly order evolutionary learning approach for working out. The category-aware design assesses the difference amid actual samples and generated samples for each class, and then uses that information to guide the model in producing high-quality category samples. Our model is further freed from difficult learning algorithms to bring up to date CatGAN on distinct statistics thanks to the Gumbel-Softmax relaxation. Furthermore, concentrating just on quality assessment typically leads to the mode collapsing problem, thus a hierarchical evolution learning technique is used to stabilize the training procedure and acquire a quality-diversity trade-off while developing CatGAN [6].

2.5 WGAN

Investigators suggest an innovative type of purpose in Wasserstein GAN that might enhance the existing GAN minimax equation. To achieve the largest probability technique work, we need to establish a noise hypothesis, nevertheless the scholars outline a arbitrary variable Z with a immovable dispersal (z) and run Z through a parametric utility that as the crow flies spawns illustrations ensuing a specific dispersal named P. So that these models could crack traditional GANs' key training issues. Training WGANs, in particular, does not necessitate finding an equilibrium in the middle of both the Discriminator and the Generator, nor does it necessitate careful network architecture design [7].

2.6 CoGAN

Coupled Generative Adversarial Networks coach couple of GANs in the identical spell. For instance, there are binary groups in a antagonism, and these double collections are comprised by twofold associates. They constitute lone appearance in binary altered fields to fuzziness the style. The Discriminator wants to separate the appearance from dual arenas. The support of an assembly be contingent on the masses they practice. Due to the collective hefts, this design essentials smaller amount of factors, associated with dual separable GANs [8].

2.7 SaGAN

Self-attention Generative Adversarial Net engenders specifics by means of indications from all characteristic positions, and it is an evolutional advancement. Old-fashioned convolutional GANs could only call the close statistics when they are skimming a picture. What's more, paralleled with outmoded convolutional GANs, the Discriminator in SAGAN is beneficial for categorize exceedingly exhaustive structures, especially in distant slices of the appearance, which are steady with one another [9].

2.8 CycleGAN

Cycle-consistent GAN resolve the problem of graphic-to-graphic transformation to produce a picture having both features from two separate species. CycleGAN is a straightforward approach of applying GANs, not a novel GAN architecture. CycleGAN acquires the plotting from solitary dominion to one more by converting unpaired images. CycleGAN, for example, may take the design from one image and transfer it into another [10].

2.9 ProGAN

Progressive GAN growth (ProGAN) aims to improve the durability of GANs in the course of the exercise stage. The influence of the Generator and Discriminator causes the training process to be unstable. This instability would occasionally result in the creation of really frightening visions. ProGAN is a training method for GNAs that keeps increasing picture quality. ProGAN first trains a 4*4 discriminator and a 4*4 generator. The produced 4*4 images which are then plotted into 16*16, 32*32, and

1024*1024, until they reach 1024 × 1024. This approach aids in the alleviation of the problem of training inconsistency [11].

2.10 StyleGAN

Style-based generator (StyleGAN) focuses on improving the generative models adequate regulator capability of the engendered appearance rather than on creating the genuine image. To do this, StyleGAN employs cutting-edge approaches such as projection networks, stylistic mixing, random variation, and others [12].

2.10.1 LapGAN

A conditional Generative Adversarial Network is the same as this. The cascade structure eases the process of each GAN's learning content, enhancing total learning ability [13].

2.10.2 AAE

Outmoded modernization fault benchmark and accusatorial working out norm are the two aims of the adversarial autoencoder (AAE). It fits the consolidated prior probability of the automatic encoder's prospective representation with any previous probability. After using a typical autoencoder to rebuild the picture as of the possible cipher, a system is accomplished to expect whether the illustration comes commencing the autoencoder's concealed code or from a subscriber sample proportion [14].

2.10.3 InfoGAN

By generative model and the common facts amid the created appearance and involvement encrypting, information maximizing Generative Adversarial Networks (InfoGAN) attempts to obtain interpretable article demonstrations via unsubstantiated learning. The advantages of this method include avoiding the use of supervised learning and huge computers to obtain easy-to-interpret features [15].

2.10.4 Invertible Conditional GAN

Invertible conditional GANs for editing images have been developed by another set of researchers [16]. They test encoders to re-generate genuine images with deterministic complex alterations by inverting the plotting of conditional GAN. When taking pictures in bad weather, such as severe rain or a blizzard, the visual quality will suffer.

The researchers seek to alter CGAN generative modelling skills via introducing an extra constriction that could eradicate the raindrop that impacts picture worth [17]. And the de-rained appearance may be blurry as of the original pristine pounded verity photograph.

2.10.5 StoryGAN

The mission of deciphering the involvement level into visuals, which may stay utilized in the direction of convey the level to bibliophiles, which is intended to be carried out by StoryGAN. The conditional generative adversarial network was used to create this model. The Context Encoder adds contextual information to the image production for each sentence. The generating process is guided by two discriminators at various stages [18].

2.10.6 ArchiGAN

ArchiGAN is a pix2pix-based program for architectural design. The workflow for this activity is separated into three parts: footmarks massing, project reallocation, and units arrangement, each having trio replicas. From the architecture design photos, ArchiGAN absorbs the topological and separate place [19].

3 Training Difficulties of GANs

GAN first surfaced in 2014, and despite the passage of six years, GAN training instability persists. Because the two neural networks diverge during the training process, GAN may not converge at all. GAN's training has been stabilized by a number of researchers. For example, biased label flatten, occasion standardization, and mini batch perception presented as solutions. This stabilization is expected to mature as GAN advances, and we will be able to train the model without issue in the near future [20]. Even when competent on numerous replica feature, GANs have the constraint of engendering mock-ups with limited miscellany. For example, when GANs are trained on data of handwritten digits with ten approaches, Generator may be impossible to create certain numbers [21]. This is known as the Helvetica scenario, and more than a few modern advancements in GANs have concentrated on resolving it.

It is also possible that instead of a fixed-point convergence, G and D fluctuate in the course of working out. When one player becomes more powerful than another, the system may fail to learn and suffer from vanishing gradient problem, resulting in variability. D quickly learns to distinguish between actual and fake samples, even when the generated samples are initially of poor quality. $D(G(z))$, the likelihood of the spawned sections being actual, will be near to 0 as a result, resulting in a very modest

gradient of log(1 $D(G(z))$ [22]. This demonstrates that G will cease updating if D fails to give gradients. In addition, choosing hyperparameters like bundle dimension, motion, weightiness decline, and mastering degree is critical for GANs training to converge [23].

We will go over the primary obstacles in GANs training in depth in this part.

3.1 Standards for Assessment

Not being constantly observed depiction erudition, watched and semi-watched erudition, in painting, denoising, and countless solicitations have all employed the GANs concept. There is a lot of variety in model construction, training, and evaluation for these broad applications. Even though numerous methodologies and processes have been recognized to weigh GAN concert, in spite of the convenience of many GANs models, the assessment is stagnant qualitative. Visual examination takes time, is subjective, and fails to capture distributional properties, which are critical in not being constantly observed erudition. The choice of proper archetypal is critical for good application concert, and the choice of proper valuation measurement is critical for reaching the correct assumption. It is necessary to construct or use appropriate quantitative metrics to overcome the constraints of qualitative measures in order to design a better GANs model. With the introduction of new models, a variety of GANs evaluation metrics have recently been provided [20].

3.2 Phase Burst

Because the max–min solution for GANs differs from the min–max solution, a phase burst hazard can develop. As a result, G* creates samples from the data distribution in $G^* = \min\theta_G \max\theta_D V(G, D)$. When $G^* = \max\theta D \min\theta_G V(G, D)$, G translates each z value to a single x coordinate, causing D to assume they are real rather than phony. Parallel gradient descent does not heavily favour min–max or vice versa. GAN oscillates and struggles to reach Nash equilibrium. Another major challenge in GANs is phase failure, which is one of the main causes of unstable GANs.

The fundamental disadvantage of GAN is they can't emphasis on the entire facts dispersal because their detached purpose is comparable to that of the JSD. Even for the binary-typical spreading, experiments have demonstrated that decreasing JSD only delivers a decent fit to the chief style and does not yield great imageries [24].

3.3 Destabilization and Non-convergence

Traditionally in GANs, *G* uses two loss functions: *Ez*[log *D*(*G*(z))] and *Ez*[log(1 − *D*(*G*(*z*)))]. Tactlessly, G's cost can clue to likely matters in GAN preparation. The prior cost task *Ez*[log*D* (*G*(z))] when *D* can easily distinguish between actual and fraudulent samples, this could be the root of the slope disappearing tricky. The minimal of the JSD amid actual and produced picture distributions for an optimal *D*, *G* loss is analogous to the minimal of the JSD amid actual and generated image distributions. As previously stated, the JSD in this scenario will be 2log2. To minimal the cross-entropy among an object category and a sorter's prophesied dissemination, the sorter must choose the accurate group. This allows optimum D to allot possibility zero to false trials and one to accurate ones, and bases slope of *G* forfeiture towards zero, which is known as fading ascents on GT. *D* aims to decrease cross-entropy while *G* attempts to enhance the same cross-entropy in GANs. *D* rejects the samples generated by *G* when *D* confidence is high, and G's gradient diminishes. Reversing the goal used to calculate the cross-entropy cost is one possible solution to this problem [20].

4 Open Issues and Future Study Areas

GANs have demonstrated their ability to generate visual features, but they still suffer from the tricky of style downfall, by way of described previously, i.e., G ruins and can only detention an inadequate number of manners in the facts. Certain of the styles qualified on the inter-modal dispersal facts are frequently forgotten by GANs. GANs rarely congregate to real symmetry and instead settle for substitute-optimum native clarifications. Because of style breakdown, the competent procreative replica's samples are generally lacking in diversity. Furthermore, we may say that style breakdown is linked to unbalanced GANs preparation, which lead to yet additional major GANs research direction.

In spite of the auspicious results, GANs remain difficult to training owed to a number of mutual unbalanced preparation and conjunction behaviours, including disappearing slopes, style breakdown, and divergent or oscillating behaviour. These difficulties in GANs training frequently obstruct further study and application in this field. Existing studies have proposed numerous techniques to address the aforementioned challenges, including developing more stable network designs, altering erudition objects, standardizing purposes, working out policies, regulation hyper parameters, and so on. Nevertheless, in most cases, achieving these goals necessitates losing appearance superiority and assortment in situations where these matters have craft-offs.

The majority of previous research has concentrated on appearance eminence or diversity. Working on appearance quality while avoiding the disadvantages of poor image diversity is one prospective study topic. Furthermore, previous strategies for

dealing with the problem of GAN training instability rely on heuristics that are extremely delicate to changes. This is one of the key motives why these approaches are limited in their applicability in new sectors. However, we notice that the majority of previous studies attempt to address one and only exercise matter at a spell and rarely involve hypothetical investigation. One more major investigation topic is to develop a theoretical agenda for dealing with difficulties in the GANs exercise procedure, by the goal of discovering additional docile inventions and making preparation more steady and simple.

Furthermore, the recommended solutions differ on the extent of training improvement. The hyperparameter settings and computational resources of most models can be improved to reach similar outcomes. As a result, we may conclude that the vast majority of similar works focused primarily on reaching up-to-the-minute accurateness rather than advanced capability. Emerging clarifications that improve on previous works algorithmically could be a forthcoming trend.

Existing efforts to dealing with GANs difficulties are centred on 3 key instructions: re-engineered net design, novel impartial roles, and optimal set of rules. The role of the objective GANs variations generally outperforms architectural GANs in terms of training, but they can't increase method assortment in the pictorial trials they generate. In these three directions, more than a few GANs designs and optimal elucidations have been offered, with suitable architecture, goal functions, and optimization strategies improving GANs training stability. Furthermore, goal functions are affected by optimization methodologies, hyperparameter medley, and the numeral of preparation stages, all of which can be investigated in imminent study for various GANs.

Nevertheless, exploration on leveraging further expertise, such as Web erudition [25], probabilistic theory variants, so on, to improve GANs training is still in its early stages. The use of a mix of suitable design, loss function, and optimization approaches can produce improved outcomes, and this could be a future research topic to pursue.

5 Summary and Conclusion

GANs have recently gotten a lot of attention for their ability to generate representative imageries, and they've grow into vital popular present day supplications including twin cohort sphere adaption, so on. GANs, on the other hand, are difficult to train, and there are three key challenges: (1) Phase Burst, (2) Destabilization and non-convergence, and (3) Standards for Assessment. Designing an effectual prototypical through picking adequate system manner, applying acceptable unprejudiced functions, or adopting appropriate optimal strategies are some of the possible answers to these GANs issues. Many different GANs versions with various characteristics have already been offered within these solutions; however, some concerns remain unaddressed.

GAN research is fairly extensive, and there are numerous designing and training techniques for GANs that can handle these issues ahead. In this post, we review the fundamental GANs architecture and look at recent breakthroughs in GAN strategy and optimal. Additional specifically, we suggested a new catalogue of GAN strategy, optimal strategies grounded on re-engineered net design, novel impartial roles, and optimal set of rules, as well as a discussion of how previous work addresses these issues. For detecting research gaps in GANs, we mapped previous works to the taxonomy. Our work serves both novices and experienced researchers by providing a snapshot of current progress as well as an in-depth analysis of the methodologies under consideration. In addition, the novel catalogue attempts to create a tricky-resolution assembly that readers can use as a guide when choosing research topics or designing techniques. We identified promising future study directions based on the information we gathered and in our future work will find solutions for GAN challenges.

References

1. Yu X (2020) IOP Conf Ser: Mater Sci Eng 740:012132
2. Wang S (2017) Generative Adversarial Networks (GAN): a gentle introduction (updated)
3. Chen D, Gao X, Xu C et al (2020) FlowGAN: a conditional generative adversarial network for flow prediction in various conditions. In: 2020 IEEE 32nd international conference on tools with artificial intelligence. 32nd International Conference on Tools with Artificial Intelligence (ICTAI 2020), 09–11 Nov 2020, Online IEEE
4. Bourou S, El Saer A, Velivassaki T-H, Voulkidis A, Zahariadis T (2021) A review of tabular data synthesis using GANs on an IDS dataset. Information 12:375. https://doi.org/10.3390/info12090375
5. Nikkath Bushra S, Shobana G (2020) Proceedings of the third International Conference on Intelligent Sustainable Systems [ICISS 2020] IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3
6. Liu Z, Wang J, Liang Z (2019) arXiv:1911.06641v2 [cs.CL]. 20 Nov 2019
7. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein gan. arXiv preprint arXiv:1701.07875
8. Liu MY, Tuzel O (2016) Coupled generative adversarial networks. In: Advances in neural information processing systems, pp 469–477
9. Zhang H, Goodfellow I, Metaxas D et al (2018) Self-attention generative adversarial networks. arXiv preprint arXiv:1805.08318
10. Zhu JY, Park T, Isola P et al (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp 2223–2232
11. Karras T, Aila T, Laine S et al (2017) Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196
12. Karras T, Laine S, Aila T (2019) A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4401–4410
13. Denton EL, Chintala S, Fergus R (2015) Deep generative image models using a laplacian pyramid of adversarial networks. In: Advances in neural information processing systems
14. Makhzani A, Shlens J, Jaitly N et al (2015) Adversarial autoencoders. arXiv preprint arXiv:1511.05644

15. Chen X, Duan Y, Houthooft R et al (2016) Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In: Advances in neural information processing systems, pp 2172–2180
16. Perarnau G, Van De Weijer J, Raducanu B et al (2016) Invertible conditional gans for image editing. arXiv preprint arXiv:1611.06355
17. Zhang H, Sindagi V, Patel VM (2019) Image de-raining using a conditional generative adversarial network. IEEE Trans Circuits Syst Video Technol
18. Li Y, Gan Z, Shen Y et al (2019) StoryGAN: a sequential conditional GAN for story visualization. In: Proceedings of the IEEE conference on computer vision and pattern recognition, p 6329
19. Stanislas C (2019) ArchiGAN: a generative stack for apartment building design
20. Park S-W, Ko J-S, Huh J-H, Kim J-C (2021) Review on generative adversarial networks: focusing on computer vision and its applications. Electronics 10:1216. https://doi.org/10.3390/electronics10101216
21. Lin Z, Fanti G, Khetan A, Oh S (2018) PacGan: the power of two samples in generative adversarial networks. In: Advances in neural information processing systems, pp 1498–1507
22. Goodfellow (2016) Tutorial: Generative adversarial networks. arXiv Preprint arXiv1701.00160
23. Radford A, Metz L, Chintala S (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. In: 4th International Conference on Learning Representations (ICLR'16)
24. Theis A, Van Den Oord, Bethge M (2016) A note on the evaluation of generative models. In: 4th International Conference on Learning Representations (ICLR'16)
25. Grnarova P, Levy KY, Lucchi A, Hofmann T, Krause A (2018) An online learning approach to generative adversarial networks. In: 6th International Conference on Learning Representations (ICLR'18)

CARSA—Smart Integrated Car Parking System

Abhinav Patil, Yash Keni, and Wricha Mishra

Abstract The increasing rate of personal automotive usage within the urban areas is a result of the aggressively growing economy, dilapidated policies, and subvention are the primary causes creating concerns for transport and automotive parking. The coordination between parking policies and traffic management is disclosed; however, parking is turning into a barrier to the through-traffic operation. Also, it is chargeable for the inefficient use of accessible resources, even the choices are created on ad-hoc basis while creating policy. The objective of this research paper is to study the parking choice of behavior made by the user and enhance the experience, making it effortless for the user to park the vehicle. This study integrates these aspects and presents the progressive review of models and studies on the parking system. The methods used to determine and research uses are contextual inquiry, data collection, flow models, surveys, usability testing using SUS and RTP. The designed HCI-based user interface, inside the car dashboard showed a substantial enhancement in parking choice of behavior made by the user and increase in environmental sustainability factor. The decreased parking finding time resulted in increased productivity of the user.

Keywords Need and capacity for parking · Features of parking · Choice of parking location · Management and planning

1 Introduction

Aside from the high rate of population expansion, faster urbanization throughout the world has resulted in higher rates of urban economy, living conditions and income, all of which contribute to the continuous rise of private automobiles. The key factors for the increase in the number of automobiles in cities, according to Shen Q, are population expansion and increased living standards [1].

A. Patil (✉) · Y. Keni · W. Mishra
MIT Art, Design and Technology University, Maharashtra, Pune, India
e-mail: abhinavp2597@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_32

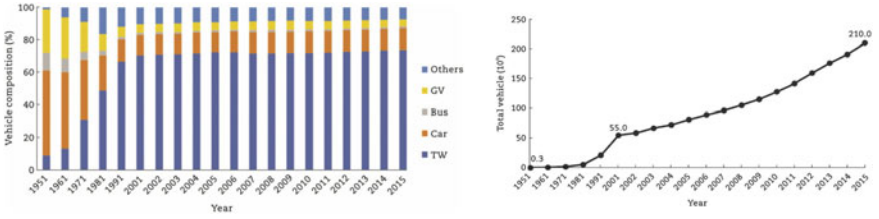


Fig. 1 Vehicle composition in India [4]

In many metropolises with short roads and uncontrolled roadside parking, the parking problem has gotten significantly more serious as a result of increasing urbanization, rising transport demand, and the expanding usage of private automobiles. Congestion issues need further care. Parking pricing and management, as well as fines for unlawful parking, can help to alleviate traffic congestion. This study aims to uncover the underlying factors that influence commuters' choice behavior patterns, which are influenced by a variety of factors such as personal characteristics and age, monthly income, gender, and other factors that have an impact on parking reservation preference, with the goal of finding solutions to improve the parking system's efficiency [2]. Through a research of people's views, we were able to identify their level of awareness about congestion and its consequences, as well as their perceptions about parking management and pricing and their effectiveness.

Most investigations included the cost of parking, the distance driven, and the time spent traveling, but these variables do not only impact parking behavior in a range of locations around the globe. In this work, a review of several research on parking features and demand projection utilizing established models over the previous few decades is presented.

A parking place is required for everyone who owns a car. As the number of automobiles increases, so does the need for parking spaces. According to the Ministry of Road Transport and Highways, India's vehicle population expanded by about 400% from 55 million in 2001 to 210 million in 2015, as seen in Fig. 1 [3]. The growth rate of motor vehicles is considerably faster in highly populated Indian cities like Delhi, Bangalore, Mumbai, and many more. Surprisingly, while the population of Delhi's National Capital Territory (NCT) grows at a pace of about 1% per year, the rate of increase in motor cars is over 7% [4].

2 Literature Review

Parking and psychological characteristics play a critical role in parking choices. For long-term sustainable growth, parking policies should be integrated with transport planning. The parking generation rate of a particular land use directly affects the parking demand. Parking policies in India are a bit more inclined toward public

transport, and private vehicles have not been considered much, but in today's scenario, we have to consider private vehicles over public transport [5].

Parking search is a complicated behavior, with travelers employing a variety of search strategies to locate a parking spot. Compared to general in-vehicle time, time spent looking for a parking spot is highly valued and can have a significant impact on drivers' parking space selection choice. Parking search time may contribute considerably to urban congestion in some situations, accounting for a large part of overall travel time. The search time impacts on a big scale as it leads to traffic congestion and also there is 0.2% delay by every individual leading to 30% traffic congestion ahead [1].

In a study in 2020, prospect theory approach was used to study Parking Space Reservation Behavior of Car Travelers from the Perspective of Bounded Rationality in Nanchang City, China. Age, monthly income, gender, and other characteristics were observed. The main factor influencing the parking reservation policy is the reservation price. In metropolitan cities, where parking spots are few, drivers typically cruise for an open parking place; the cruising procedure takes a long time and can result in excessive fuel consumption, harmful pollutants, and even car accidents. Drivers tend to spend less than the optimal reservation price. As the reference point for highest prices grows higher, and they become more sensitive to losses [6].

In metropolis cities, GIS technology can be used to better analyze traffic and parking, as well as identify trends, make data comparisons, and develop action plans to help mitigate parking problems and generate management planning solutions for traffic operation effectiveness. It was found out that 20–30% of the public, and underground parking lots stay vacant, while the curb side parking spaces get occupied by 120–150% [2].

In 2003, level of acceptability of various urban transport pricing strategies was examined by surveying 952 motorists in four European cities: Athens, Como, Dresden, and Oslo. The authors tested the effects of time-differentiated cordon pricing, increase in parking charges and fuel taxes, and revenue hypothecation [7].

In a 2018 study, perceptions of people in Patna, Bihar, India, were studied using structured questionnaires. A congestion pricing strategy was proposed by examining the main factors and narrowing down the underlying factors determining level of awareness of people about congestion and its effects, their willingness to take responsibility, their perception of parking management and pricing, and its effectiveness. It was concluded that illegal parking penalty can help in improving the congestion scenario of the city [8].

In 2018, an algorithm was proposed to identify the most efficient route between user and parking place based on the real-time traffic. The proposed algorithm will help users to identify parking places with minimum cost. The average waiting time for locating a parking spot has been lowered by the proposed system. Using the sensor positioned at every parking spot, this system informs users about real-time parking availability. This technology also lowers the cost of parking a car in a designated parking spot by determining the most effective path to the parking spot based on real-time traffic congestion in the area [9].

Installing a manual or digital board at the entrance of a parking lot that provides information about the status of parking spaces and the number of available empty parking spaces and will help reduce search time and encourage the drivers to use off-street parking. Increased fees for on-street parking will encourage long-term parkers to use off-street parking. The lack of suitable parking markings led to haphazard parking at both parking facilities, resulting in inefficient use of available parking space. In the case of curb parking, it encourages cars to park illegally, thus reducing the effective road width and traffic flow speed [10].

3 Methodology

3.1 Design Study and Participants

A survey was conducted with 88 working professionals and five parking lot owners and customers in the city of Mumbai and Pune in the month of October–November 2021. These factors include daily commuting, commuting experience, parking facilities near the workplace, time spent in finding a parking lot, parking lot preference, and parking assistance. Exclusion factors include time taken to reach the destination after the vehicle has been parked.

3.2 Data Collection

A total of 88 subjects were selected based on criterion sampling method, and they participated to take part in this research (age group from 16 to 51+ years). For selecting subjects, the following criterion was used:

a) Subjects commuting daily to work by car.
b) Subjects staying in Mumbai and Pune city.

When selecting and dividing individuals into groups depending on their gender, further, sampling was done on occupation (working professional, business, students, housewife, and consultant), gender, marital status, and presence of kids or pets.

3.3 Contextual Inquiry

This approach to data collection gave an understanding of how drivers make decisions and behave in their natural surroundings by observing them and interviewing a representative sample of them while they were parking, providing a comprehensive view of how drivers make decisions and park in their natural environments [11].

3.4 Competitive Analysis

The current and potential competitors in car parking were identified and evaluated on the basis of primary, secondary features, and human factors. Parking is a multifaceted issue, and this method has been used to identify existing solutions in Maharashtra. The competitors were mainly YourParkingSpace and Park + .

3.5 Brainstorming and Conceptualization

A focus group of nine drivers brainstormed concepts to solve urban car parking issues. Ideas including shared parking, technological interventions, updating norms, parking prices, mobile apps, RFID tag signals, application, filters, security system, etc., were prioritized with affinity mapping, which gave weight to the features and concepts. The final concept was finalized and designed.

4 Result

4.1 Competitive Analysis

We made the competitive feature analysis chart of the existing mobile app-based platforms namely YourParkSpace, Parking, ParkingRhino, and Park+ (Fig. 2).

4.2 Contextual Inquiry

See Figs. 3 and 4.

4.3 Data Collection

With the help of the questionnaire, the users were asked about their daily commuting experience and what time of the day they face the problem. Along with this, a thorough research was conducted for the cause of parking choice of behavior made by the user. The main probable cause for traffic jams was the inconsiderate parking and unnecessary space taken by the car owners while parking the vehicle (Figs. 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14).

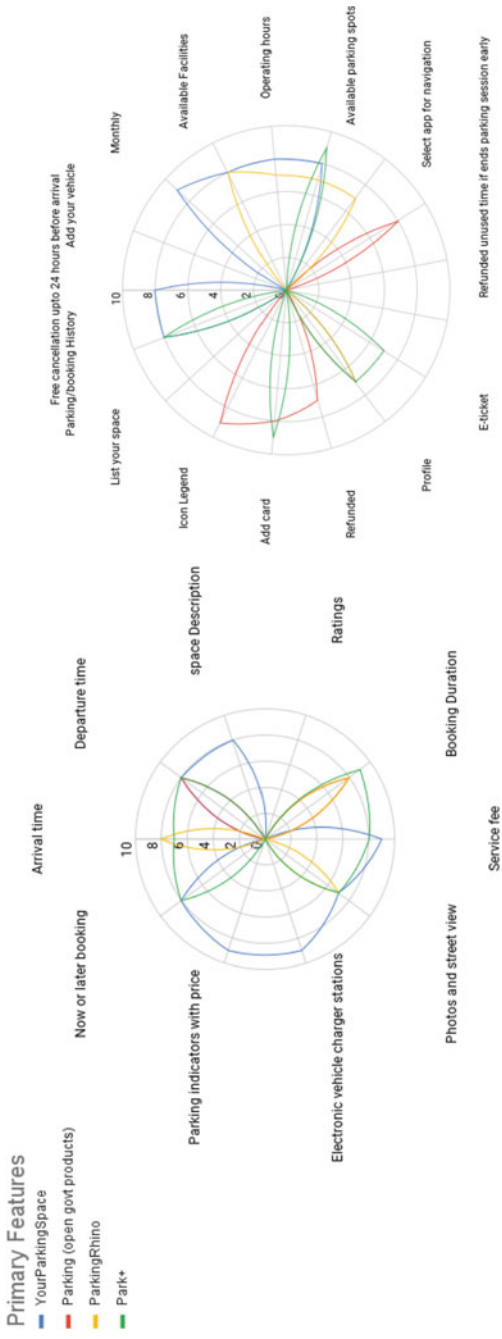


Fig. 2 Competitive feature analysis

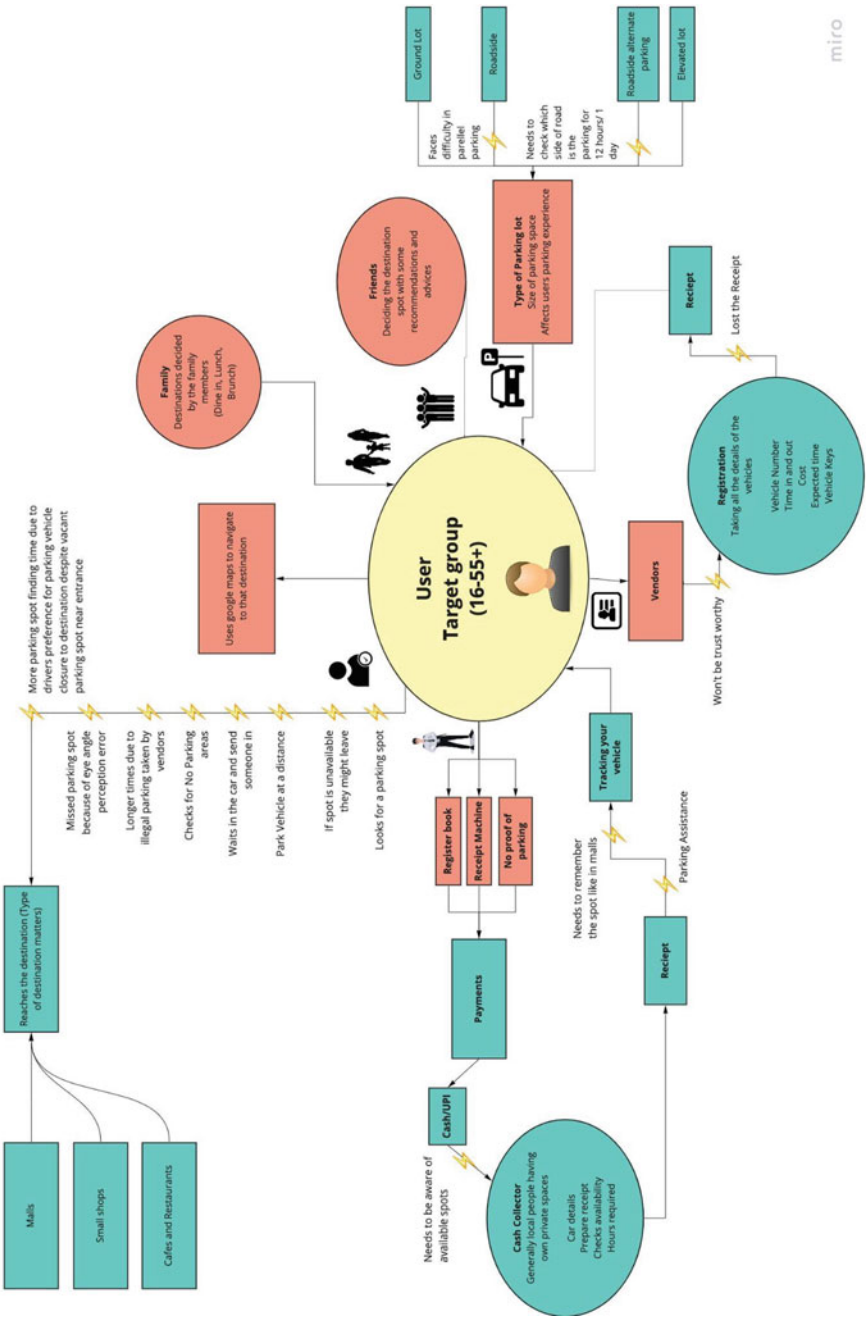


Fig. 3 Flow model

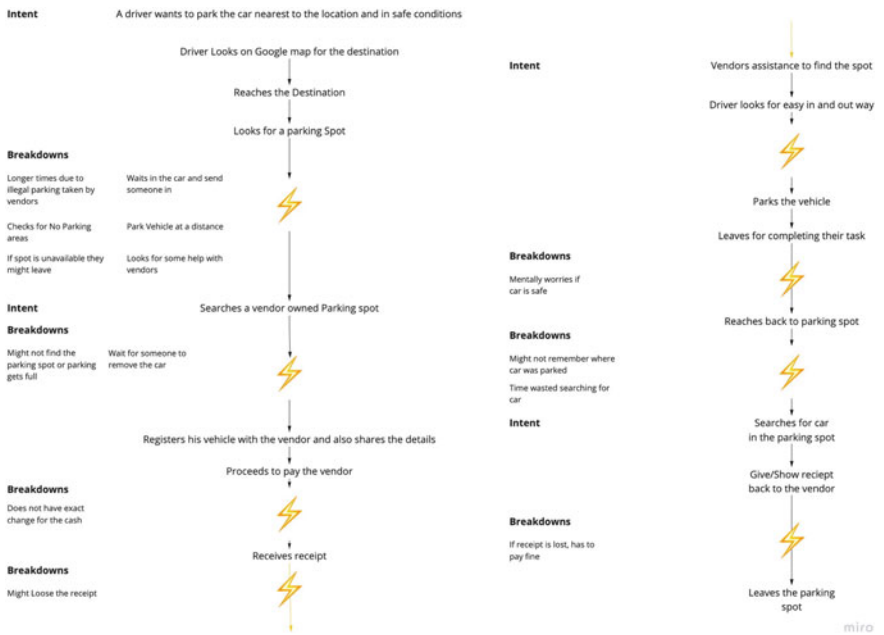


Fig. 4 Sequence model-driver wants to park the car nearest to the location and in safe condition

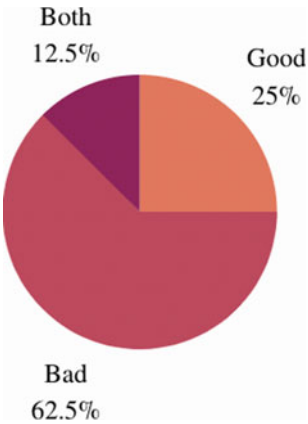


Fig. 5 Users daily commuting experience

5 Design Intervention

The upcoming technology is focusing on more user-friendly and easy accessibility to car dashboards. The competitors in the market are looking forward to integrating as many features as possible along with solving maximum daily user problems in the Indian scenario.

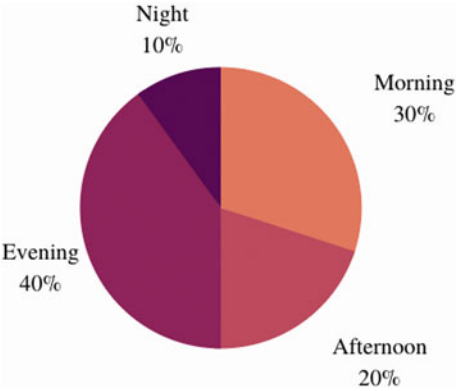


Fig. 6 Time of the day user face this issue?

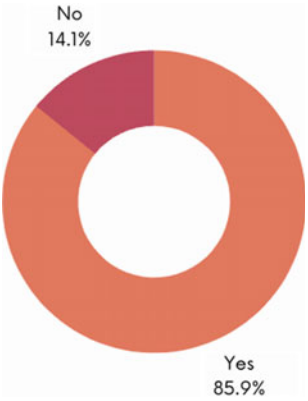


Fig. 7 Do users find vehicles are parked obstructively, unfairly, and/or inconsiderately in your road

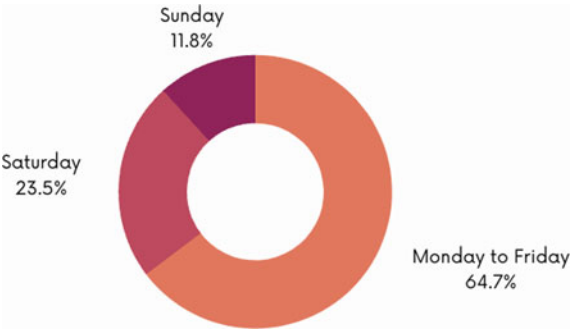


Fig. 8 What time of the week do they face this issue

Additionally, the automobile industry is slowly transforming toward more development and use of EV vehicles which have more dependency on dashboards and ADAS technology. More government initiatives pertaining to EV's.

Fig. 9 Why do they face this issue?

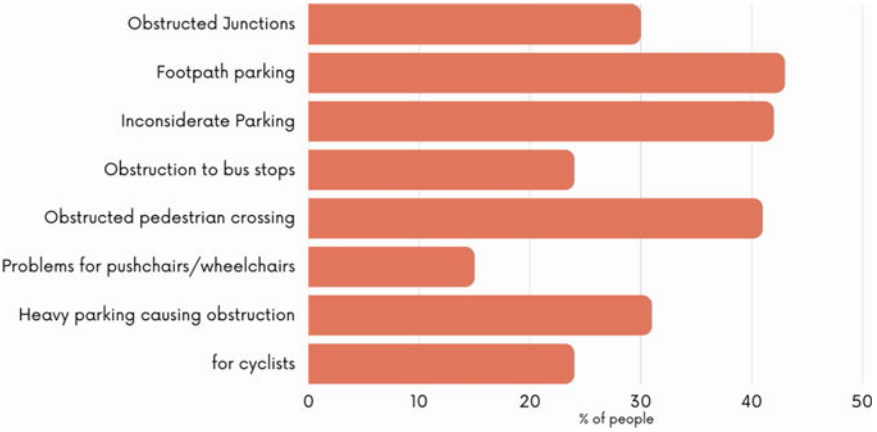
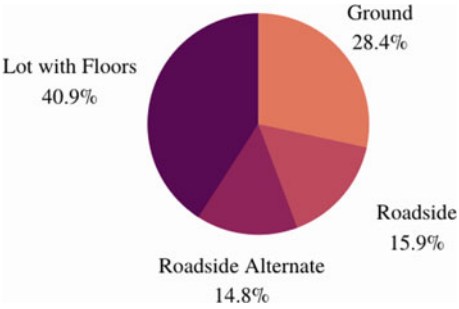


Fig. 10 Do parked vehicle in your road cause you problem as a pedestrian or non-motorist (i.e., as a cyclist, etc.)

Fig. 11 What kind of parking lot would you rather prefer?



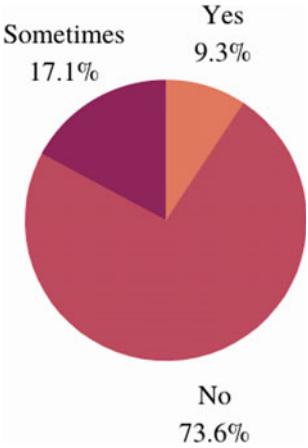


Fig. 12 Do you receive parking assistance?

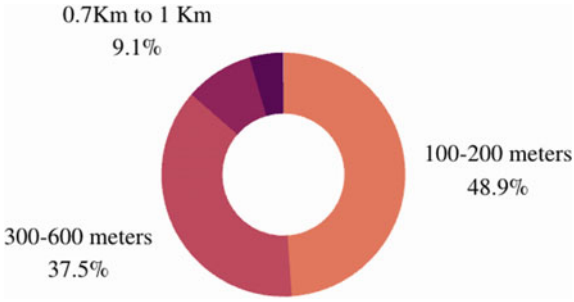


Fig. 13 How often do you travel on weekends?

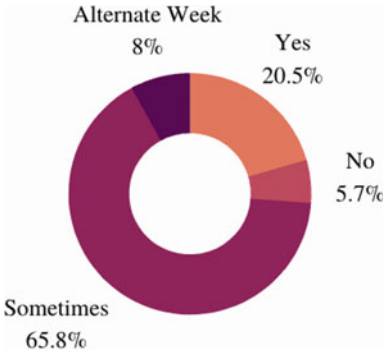


Fig. 14 How far from your destination are you willing to park your car

In order to reduce parking spot searching time and make the process of booking parking spots easier, convenient, and more reliable, we propose a solution "CARSA" that would help car drivers effectively search and book parking spots through the car dashboard system (also called dash, metal panel (IP), or fascia). It is a control panel placed inside the center console of a car. It is usually available directly before the

Fig. 15 User interface design screens

driver and displays a large amount of information based on the model purchased by the user.

Many cars today offer many advanced features at a much lower price. A few of them are the Mahindra XUV700 launched in September 2021. This car offers the new ADAS as shown in Fig. 16 also known as advanced driver control. The main function of ADAS is to guide the driver and identify any obstacles along the way.

This ADAS feature can also help the car dashboard display the available parking spot near his/her destination which helps the user select a parking choice based on the choice of behavior. Few pros of this concept are flexibility using the system, adaptable by the user, it can be cost efficient as it gets integrated or sits as an additional feature to the car, reduces parking time spent by the user and provides safety with plenty of features with multi-platform syncing. But for a few additional features a certain amount of expenses will be incurred. Hence, a parking assistance system inside the car will be a great addition to solve a lot of problems for the user (Figs. 15 and 16).

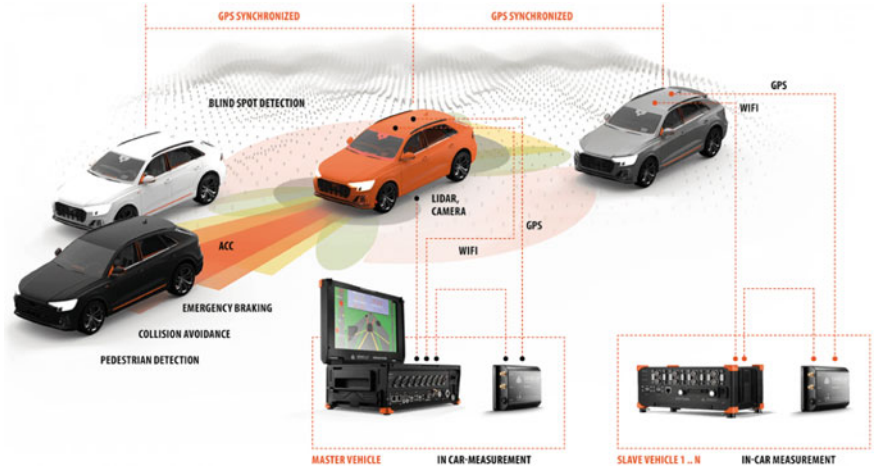


Fig. 16 ADAS system [12]

6 Usability Testing

In order to gain insight into the acceptance of the design concept, a few car drivers were interviewed. It was found that the drivers really loved the idea and wanted to see it developed further. As a result, they discussed how the issues and difficulties would have been improved if they had been able to experience it properly with the help of smart parking dashboard integration. The participants found it user-friendly and intuitive. SUS and RTP tests were conducted to determine usability. According to our analysis, the design and usability performance of the design was rated excellent on average by the users with a SUS score of 81 (Figs. 17 and 18).

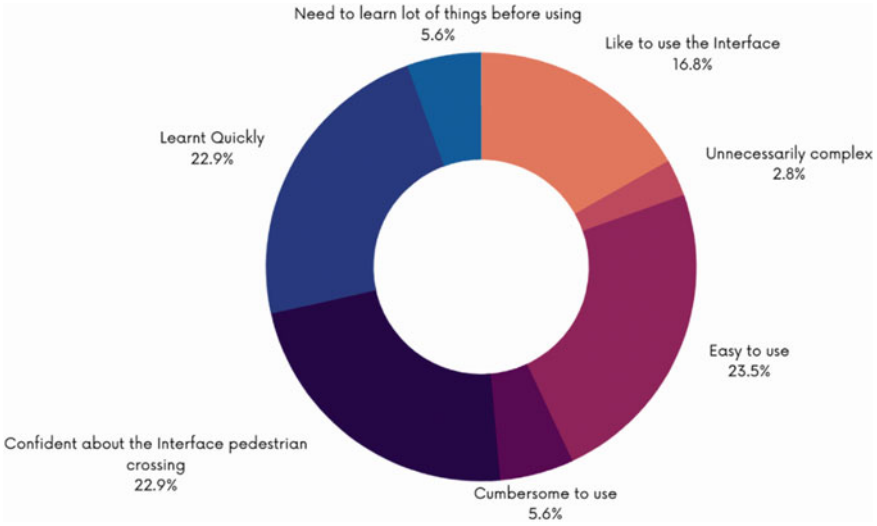


Fig. 17 Usability testing using system usability scale (SUS)

Tasks (Sample size = 20)	Overall Effectiveness (%)	Time based efficiency (goals/sec)
Task 1 (Book a parking spot - start to end process)	100	0.0431
Task 2 (Change booking time slot to 5 hours 40 mins)	80	0.0748
Task 3 (Click on book now in parking space description)	100	0.0824

Fig. 18 Usability metrics

7 Discussion

This study analyzed the problems faced by car drivers searching parking space in Mumbai and Pune city, and it helped gather insights specifically in terms of parking spot search behavior. The questionnaire survey's data processing analysis confirmed that a traveler's age, gender, monthly income, and other parameters impact their parking reservation selections. The reservation price is the most important component in determining parking reservation policy [6]. Findings from a self-reported questionnaire also showed that 69% car drivers found difficulty finding parking space everyday while commuting many drivers. 85.9% people find vehicles parked obstructively, unfairly and inconsiderably on the road. 64.7% people find difficulty in finding parking spots on weekdays Monday to Friday followed by 60.2% in evening time being most busy. 40.9% of drivers preferred parking lots with floors followed by 28.4% ground lot parking. Thus, a need for a smart and personalized impactful parking spot finding system was identified for car drivers to make them more stress free about the parking aspect of the journey. Additionally, given the difficulties in digitizing multiple unorganized parking spots, CARSA can help construct a collective environment for both organized and unorganized parking lot owners with an easy onboarding and registration process for both the parking lot owners and car drivers.

The advantages of our parking research include the utilization in potential use of the omnipresent dashboards in cars, and eliminating the distraction caused by attaching a mobile to the glass window shield while driving a car. Thus, reducing possible risks of accidents while navigating and cruising while searching for parking spots.

8 Limitations

The scenarios might change city to city as the study was performed only in two cities, i.e., Mumbai and Pune. Apart from this, the ADAS system needs a digital touch instrument system to detect various obstacles which all the cars will not have. But the scenario in the future might change where every car model will definitely have an infotainment system.

9 Conclusion

Considering the rise of personal car vehicle growth in India as well as the problems generated by the in this research, we discovered various breakdowns that the drivers face while parking their vehicles at their destination by conducting a survey of 88 participants from an age group from 16 to above following the necessary research methodologies. We learned how various factors like age, occupation, marital status,

time and day of travel, and pricing plays an important role toward parking behavior of drivers. Final concept of "car dashboard/screen integrated AI-based software" was chosen for further development. Based on the usability testing, we found out that the increased efficiency of parking a spot from the current destination helps the user save a lot of time and also reduces the carbon footprint which eventually leads to a sustainable future.

10 Future Scope

For future work, we intend to expand on this research by prototyping and integrating the smart dashboard software in cars and testing it in Maharashtra to help resolve the present parking spot searching problem. The study can be further extended to help develop more parking system design solutions from the insights gathered in this research. We plan to extensively explore the key approach for more parking behavior factors since there are more cognitive factors that can be studied to give more insights to the parking problem and to provide a holistic view.

References

1. Axhausen KW, Polak JW (1991) Transportation planning: activity-based approach. In: Concise encyclopedia of traffic & transportation systems. Elsevier, pp 564–569
2. Liu Y, Wang W, Ding C, Guo H, Guo W, Yao L, Xiong H, Tan H (2012) Metropolis parking problems and management planning solutions for traffic operation effectiveness. Math Probl Eng 2012:1–6. https://doi.org/10.1155/2012/678952
3. Ministry of Road Transport and Highways. https://morth.nic.in/. Last accessed on 16 Mar 2022
4. Home, http://desagri.gov.in/. Last accessed on 16 Mar 2022
5. Parmar J, Das P, Dave SM (2020) Study on demand and characteristics of parking system in urban areas: a review. J Traffic Transp Eng (Eng Edn) 7:111–124. https://doi.org/10.1016/j.jtte.2019.09.003
6. Xue Y, Cheng L, Lin P, An J, Guan H (2020) Parking space reservation behavior of car travelers from the perspective of bounded rationality: a case study of Nanchang City, China. J Adv Transp 2020:1–13. https://doi.org/10.1155/2020/8851372
7. Schade J, Schlag B (2003) Acceptability of transport pricing strategies. In: Acceptability of transport pricing strategies. Elsevier, pp 1–9
8. Bandyopadhyaya V, Bandyopadhyaya R, Prashad C (2018) A study on awareness and effectiveness of congestion management through parking pricing for Patna, India. In: Proceedings of the 40th international academic conference, Stockholm. International Institute of Social and Economic Sciences
9. Zajam A, Dholay S (2018) Detecting efficient parking space using smart parking. In: 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE
10. Parmar J, Das P, Azad F, Dave S, Kumar R (2020) Evaluation of parking characteristics: a case study of Delhi. Transp Res Proc 48:2744–2756. https://doi.org/10.1016/j.trpro.2020.08.242
11. Peyrot M, Darroch V, Silvers RJ (1983) Interpretive human studies: an introduction to phenomenological research. Contemp Sociol 12:561. https://doi.org/10.2307/2068731

12. ADAS testing—advanced driver assistance systems testing with Dewesoft, https://dewesoft.com/applications/vehicle-testing/adas. Last accessed on 16 Mar 2022

An Approach to Extract Major Parameters of Legal Documents Using Text Analytics

Souraneel Mandal, Sajib Saha, and Tanaya Das

Abstract Natural language processing is one of the fascinating areas of artificial intelligence to comprehend human language. It has been widely used to address various issues to understand human language by computers. It has been utilized to pre-process unstructured data and investigate authoritative records to help in legitimate decision-making. The majority of data found in the Indian legal system are unstructured. Various courts disperse legal documents by genuinely dissecting the data of particular cases and getting the choice from perspectives decisions and the rule regulation. Legal professionals' traditional way of analysing, and making decisions is time-consuming. Hence, in this paper, the authors have focused to fabricate a text pre-handling procedure on authoritative reports to eliminate undesirable texts. The authors have proposed a methodology to pre-process legal documents for content characterization, and text arrangement and examine their specialized commitments to find the major words found in legal documents. Natura language processing techniques like bag of words, count vectorization, etc., help to examine the catchphrases of the court procedures after removing the unwanted texts from legal documents to extract major words found in legal documents.

Keywords Dowry death cases · Natural language processing · Legal documents · Lemmatization · Count vectorization

1 Introduction

In the twenty-first century, technology is permeating many facets of our lives. Artificial intelligence (AI) is one of the sophisticated technologies that enable us to solve a wide range of real-world concerns such as decision-making, prediction, extracting meaning from imprecise or raw data to identify patterns, recognizing trends that are too complex for the human brain for last ten years, adopting and implementing technology that helps to accommodate and make life easier. To accomplish the artificial

S. Mandal (✉) · S. Saha · T. Das
Adamas University, Kolkata, India
e-mail: souraneel97@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_33

intelligence framework, there are popular domains are available in the industries like machine learning [1] which empowers computers or machines to create data-driven choices instead of being expressly modified for a certain errand, fuzzy logic [2] which approach is rationale mimics the way of choice-making in people that includes all middle conceivable outcomes between advanced values yes or no, natural language processing (NLP) [3] which utilized by machines to get it, dissect, control, and translate human's dialects which makes difference engineers organize information for performing assignments such as interpretation, programmed summarization, discourse acknowledgement, and relationship extraction.

NLP is using in various sectors for text pre-processing like spam filters [4], spell check [5], autocomplete [6], application in legal system [7], etc. Indian legal system has many spaces where NLP can help in various ways as most of the legal documents are available in unstructured format. Use of NLP can help to reduce pending litigations, infrastructure deficiencies that results delay in decision-making. Currently, most advocates manually collect data and compare it to previous judicial pronouncements and legislation before presenting the final decision to the court. Legal professionals still use traditional methods of legal research to extract the major parameters from legal documents. They manually go through the documents which is time-consuming in nature. Throughout India, several incidents of one-sided rulings have been documented. Furthermore, the choice will be influenced by the political parties in power, making it frequently erroneous and illegal. Hence, the authors in this paper have proposed an approach to pre-process unstructured legal documents and extract the major parameters present in legal cases. By utilizing this approach of NLP techniques like regular expression, bag of words, rhetorical function sentence labelling, etc., will help legal professionals to anticipate proper and transparent verdicts in accordance with the law. Rhetorical function sentence labelling is the process of determining what semantic function a sentence in a legal document is related to, such as case facts, parties' arguments, and the court's ultimate judgement. Identifying the rhetorical purposes of terms in a legal case document can help with things like semantic search, summarization, and case law analysis. It would be really beneficial for legal professionals to comprehend the expected outcome in his case. The computerized examination of the legal facts provided will encourage many courts to publish them online, making them available to the public for future reference. This study looks at the possibility of utilizing natural language processing and neural networks to improve statistical research in the legal areas. The possibility of natural language processing techniques for autonomously forecasting court judgments in Indian courts is established and thoroughly studied.

Section 2 describes the related study on natural language processing techniques used for legal system. Section 3 illustrates the proposed methodology to extract the major words available in legal cases. Finally, Sect. 4 concludes our proposed work along with its future utilization in Indian legal system.

2 Literature Review

Nowadays, large volume of data is available in both structured and unstructured from. Mostly, unstructured data are available in textual form. Hence, recently, the research community has given emphasis to the text pre-processing techniques to extract meaning information [8]. In NLP, there are various tools and technique that help to pre-process large amount of text documents, for example, word embeddings, stop words removal, etc.

In the early year 1999, some research work had been done on neural network using adaptive learning technique for information collection in legal text documents. Early, in 1999, few researchers had done research on neural network using adaptive learning techniques for information collection in legal text documents. They have labelled the neural processing elements with the most adequate keywords of the document to describe the context. Apart from successful classification, self-organizing maps are a promising method for this [9]. Same year, some researchers also proposed that discourse analysis produces useful information that may be included into text processing systems. Discourse pattern knowledge has already been used in legal text creation systems. However, incorporating this type of knowledge into legal information extraction systems is also critical. This expertise comes in handy while looking for information in texts. They have also demonstrated adequate, adjustable, and clear sharable knowledge representation of discourse patterns in the research [10]. In the year 2014, an author described the data may be highly structure or unstructured or semi structured in the documents. So, there is a big problem in retrieving information. Because the huge number of information requires more intelligent machines. For improved grouping, the research provides a novel approach for grouping case notes (abstracts) of legal document datasets [11]. After few years, a model was proposed in 2017 to pre-process text using word embedding techniques in a dynamic way and have described a system leveraging semantic models in a smart review environment intending to support knowledge workers in information gathering [12]. In the year 2018, an article has been published on systematical and empirical manner on legal juridical language which is called corpus analysis and introduced a computer-assisted legal linguistic method which will emphasize and produce a more sophisticated output from those legal documents. With the help of those corpus lawyers can evaluate legal semantics, language, and socio semiotics in various working circumstances using computer-assisted analysis of well pre-processed collections of legal texts including judiciary, legislature, legal academia [13]. Recently, LegalAI was introduced in 2020 which proposed future and present scope of research using NLP. It has described LegalAI's history, current situation, and future research goals. They have also discussed the tasks from the perspectives of legal practitioners and NLP academics, with examples in LegalAI. To explore alternative future approaches, the authors had performed tests and provided a detailed analysis of the benefits and drawbacks of current activities [14].

A variety of supervised and unsupervised machine learning and natural language processing approaches such as topic modelling, word embeddings, and transfer

learning have been reviewed in recent year 2021, that can be used to investigate legal texts on a wide scale using unsupervised learning techniques like word cloud, latent semantic analysis, etc. And also, some supervised learning like bag of words method, transfer learning, and transformers are used to obtain labelled documents [15].

From overall literature, the authors have found that there is a gap in the pre-processing techniques that has been not described in any of the literature. Hence, in this paper, the authors have focused to produce a methodology to pre-process legal documents to do further analysis on the legal documents whether it is for information processing or analysis purpose are described in further section.

3 Proposed Work

The proposed methodology includes bag of words, regular expression, natural language processing (NLP) techniques like tokenization like sent_tokenize, regex, tokenization methods, stop words removal, lemmatization like WordNetLemmatizer, pdftotext modules, and so on are included in the suggested methodology. Furthermore, the words extracted using text analytics methods such as TF-IDF, parts-of-speech tagging can be represented as the entities, their relationships, and their occurrences to give legal professionals an ontology [16]-based comprehensive solution to analyse legal cases like dowry death, domestic violence [17], etc. (Fig. 1).

Step I explains the text authors retrieved from all dowry death legal documents using the glob module, where all the pdf names are preserved as a list using a regular expression. The data(s) for each pdf document are collected page by page using the pdftotext module, and we have stored all of the pages as a string in a list and merged the list in a string. After the tokenizer has been set up, **Step II** discusses how to put tokenization into action. After that, the full string was tokenized. After the stop words were eliminated, the tokens were lowercased. The lemmatizer is then applied to each token individually. **Step III** outlines the cleaning of all unnecessary and undesired strings, with the exception of tokens containing no more than two alphabets. We have now eliminated the term that contains digits in the form of letters. After filtering, we removed all of the last names from the string lists. **Step IV** describes the CountVectorizar function to find highest frequent words in ascending order represented as DataFrames in a list. **Step V** illustrates how we created a single DataFrame that contains the number of times each phrase appears in all legal documents on dowry death using an outer join on each DataFrames in the list. **Step VI**

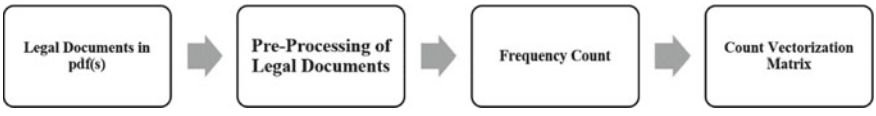


Fig. 1 Stepwise proposed methodology

```
START
1. Import nltk, glob, pdftotext, collections, pandas, beautifulsoup4 modules

2. Data: st ← sorted order(glob(path specified documents)), max_count ← 500,
        N ← length(st) and ls ← NULL, text ← " " .
3. Result: merged_df

4. for i = st[0] to st[N] do
       f ← open i^th document in read mode
          pdf ← pdftotext.PDF function(f)

       for j = 0 to length(pdf) do
          Append j^th page as string inside ls
       End

       text ← join all the list
       word_tokens ← Lower Case (RegexpTokenizer(text))
       Initialize lemmatizer ← WordNetLemmatizer(text)

       word_tokens ← append(lemmatizer.lemmatize(word_tokens))
       word_tokens ← append(len(word_tokens) > 2)
       word_tokens ← remove(word_tokens.isdigit())
       word_tokens ← remove(BeautifulSoup(word_tokens).get_text())
       word_tokens ← remove(stopwords.words('english'))
       word_count ← collections.Counter(word_tokens)
       Max_MostFreq ← word_counts.most_common(max_count)
    End

5. merged_df ← apply Outer-Join(Import Max_MostFreq to DataFrame)
6. merged_df ← replace NaN with 0
7. merged_df
END
```

Fig. 2 Proposed pseudocode

shows how to convert a DataFrame from category to integer columns, maintaining only the rows with at least two non-null values and replacing all null values with 0.

Based on the proposed methodology, the authors have proposed a pseudocode as shown in Fig. 2.

4 Result and Discussion

As shown in Fig. 2, the pseudocode of our proposed work successfully indicates the major words present in legal documents that are extracted by the legal professionals to prepare their draft. Hence, in Fig. 3, the authors have prepared a matrix which shows the count of each word frequency present in any legal documents.

In Fig. 4, the authors have extracted top 50 words with highest word count frequency across legal documents that have been taken and visualized as a bar plot.

	word	Case1	Case2	Case3	Case4	Case5	Case6	Case7	Case8
0	person	65.0	22.0	1.0	39.0	8.0	10.0	10.0	0.0
1	accused	56.0	15.0	4.0	35.0	21.0	9.0	33.0	8.0
2	evidence	49.0	15.0	37.0	34.0	45.0	44.0	9.0	6.0
3	husband	46.0	16.0	2.0	27.0	19.0	12.0	8.0	4.0
4	dowry	46.0	12.0	16.0	28.0	12.0	21.0	14.0	16.0
...	...	...	...	...	...	...	...	...	...
127	mother	1.0	0.0	11.0	15.0	12.0	16.0	6.0	4.0
128	examined	1.0	8.0	3.0	9.0	2.0	0.0	4.0	1.0
129	place	0.0	8.0	1.0	14.0	8.0	9.0	2.0	6.0
130	told	0.0	7.0	6.0	8.0	11.0	4.0	2.0	1.0
131	could	0.0	6.0	5.0	3.0	5.0	3.0	5.0	12.0

Fig. 3 Pseudocode for text pre-processing for legal documents

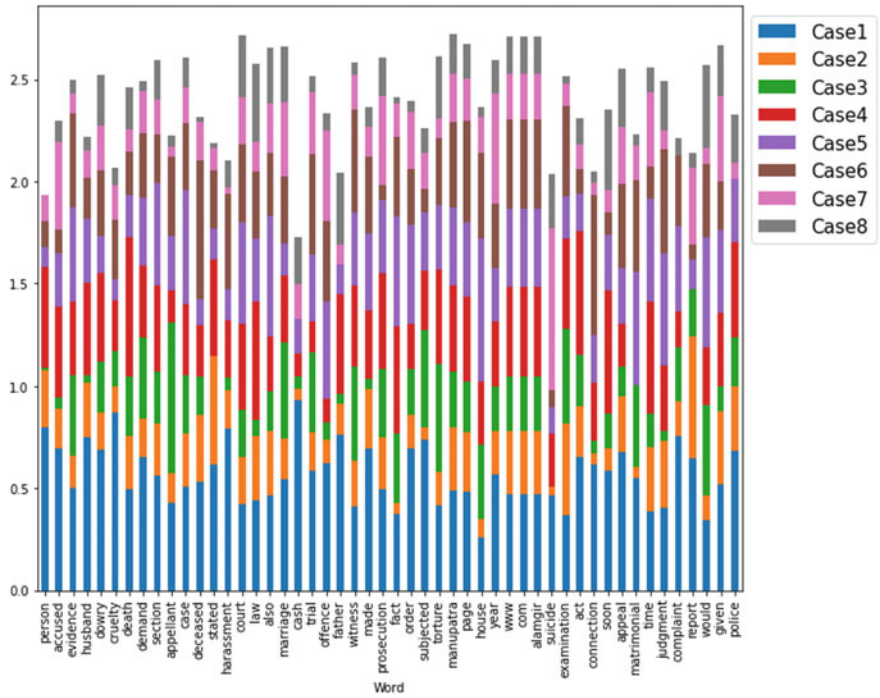


Fig. 4 Visualization of top 50 frequent words

5 Conclusion and Future Work

Text pre-processing of legal documents plays a vital role for widening the view of enabling tools to help the legal professionals to extract essential information hidden in legal documents. This proposed methodology extracts the essential words found in legal documents that helps legal professionals to prepare their draft in less time. The authors have proposed an algorithm to pre-process legal documents using natural language processing (NLP) techniques like count vectorization method to find most frequent words represent in legal documents. These frequent words help legal professionals to understand the major parameters found in legal documents. In future, these words can be further represented as the statistical summary of the legal documents using concepts like TF-IDF, N-gram, etc.

References

1. Trappenberg TP (2019) Machine learning with sklearn. Fundamentals Mach Learn 38–65. https://doi.org/10.1093/oso/9780198828044.003.0003
2. Fuzzy Logic Theory (2000) Introduction to fuzzy sets, fuzzy logic, and fuzzy control systems. https://doi.org/10.1201/9781420039818.ch2
3. Mochihashi D (2020) Robotics, grounding and natural language processing. J Natural Language Process 27(4):963–968. https://doi.org/10.5715/jnlp.27.963
4. (2014) An improved spam filter for filtering repeated spam e-mails. Int J Adv Eng Res Dev 1(05). https://doi.org/10.21090/ijaerd.01056
5. Mati DN, Hamiti M, Selimi B, Ajdari J (2021) Building spell-check dictionary for low-resource language by comparing word usage. In: 2021 44th international convention on information, communication and electronic technology (MIPRO). https://doi.org/10.23919/mipro52101.2021.9597183
6. Voytovich L, Greenberg C (2021) Natural language processing: practical applications in medicine and investigation of contextual autocomplete. Acta Neurochir Suppl 207–214. https://doi.org/10.1007/978-3-030-85292-4_24
7. Frankenreiter J, Nyarko J (2022) Natural language processing in legal tech. SSRN Electron J. https://doi.org/10.2139/ssrn.4027030
8. Raj A, D'Souza R (2021) Anonymization of sensitive data in unstructured documents using NLP. Int J Mech Eng Technol (IJMET) 12(4). https://doi.org/10.34218/ijmet.12.4.2021.002
9. Schweighofer E, Merkl D (1999) A learning technique for legal document analysis. In: Proceedings of the seventh International Conference on Artificial Intelligence and Law—ICAIL '99. https://doi.org/10.1145/323706.323790
10. Moens M-F, Uyttendaele C, Dumortier JOS (1999) Information extraction from legal texts: the potential of discourse analysis. Int J Human-Comput Stud 51(6):1155–1171. https://doi.org/10.1006/ijhc.1999.0296
11. Wagh RS (2014) Exploratory analysis of legal documents using unsupervised text mining techniques. Int J Eng Res Technol. Retrieved April 14, 2022, from https://www.ijert.org/exploratory-analysis-of-legal-documents-using-unsupervised-text-mining-techniques
12. Vo NP, Privault C, Guillot F (2017) Experimenting word embeddings in assisting legal review. In: Proceedings of the 16th edition of the international conference on artificial intelligence and law. https://doi.org/10.1145/3086512.3086531
13. Vogel F, Hamann H, Gauer I (2018) Computer-assisted legal linguistics: corpus analysis as a new tool for legal studies. Law Soc Inq 43(04):1340–1363. https://doi.org/10.1111/lsi.12305

14. Zhong H, Xiao C, Tu C, Zhang T, Liu Z, Sun M (2020) How does NLP benefit legal system: a summary of legal artificial intelligence. In: Proceedings of the 58th annual meeting of the association for computational linguistics. https://doi.org/10.18653/v1/2020.acl-main.466
15. Dyevre A (2021) Text-mining for lawyers: how machine learning techniques can advance our understanding of legal discourse. Erasmus Law Rev 14(1). https://doi.org/10.5553/elr.000191
16. Das T, Roy A, Majumdar AK (2020) A study on legal knowledge base creation using artificial intelligence and ontology. Innov Data Commun Technol Appl 647–653. https://doi.org/10.1007/978-3-030-38040-3_73
17. Das T, Roy A, Majumdar AK (2022) Ontology based knowledge visualization for domestic violence cases. Innov Bio-Inspired Comput Appl 579–589. https://doi.org/10.1007/978-3-030-96299-9_55

AR/VR-Based Comprehensive Framework for Virtual Convocation

Nilay Udeshi, Soham Dhuri, Diya Luniya, and Neha Deshmukh

Abstract In this pandemic, people have not been able to celebrate their achievements and accomplishments with their friends and family. One of those achievements is the convocation/graduation ceremony of the students who have successfully invested their hard-work and dedication for achieving this prestigious milestone. The number of people that can attend a physical graduation ceremony is limited due to restricted venue space, seating, and food. Students are unable to invite all of their family members and friends to their big day. Students can invite as many family members as they like to a virtual convocation event. Their family may be anywhere in the world, but they would not miss this momentous occasion. No one is left behind, whether it's kids, parents, or their family and friends, thanks to live transmission to remote audiences and on-demand post-event sessions. Students often return to their native country and are unable to attend their graduation ceremonies. This enables them to attend and create memories with their friends, classmates, and family. This is where the project idea comes into play, where a cross-platform app will be made that will be able to provide a real-life environment to the students and their family members where they will be able to enjoy the convocation ceremony without breaking any social distancing norms.

Keywords Unity3D · 3D models · Virtual environment

N. Udeshi (✉) · S. Dhuri · D. Luniya · N. Deshmukh
Department of Information Technology, A.P. Shah Institute of Technology, Thane, MH, India
e-mail: nilayudeshi@apsit.edu.in

S. Dhuri
e-mail: sohamdhuri@apsit.edu.in

D. Luniya
e-mail: diyaluniya@apsit.edu.in

N. Deshmukh
e-mail: npdeshmukh@apsit.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_34

1 Introduction

Virtual platforms like Zoom, Cisco Web-Ex, and Microsoft teams have proved to be very useful for conducting meetings and conferences during this pandemic situation as people are not able to meet each other in person for such occasions. The graduation ceremony is not available to many online learners, and thus, they cannot commemorate this important milestone. It starts by assessing the literature on virtual graduations and creating a typology of several graduation models used by higher education institutions today. As a solution, the idea aims to develop an interactive system that will help colleges to conduct a convocation ceremony for their students in an orderly and effective way.

A virtual graduation ceremony is one that takes place over an online event platform. It offers a virtual environment that is similar to a real-life commencement event where students get their academic diplomas or degrees. If the virtual atmosphere is accurate, it can bring your event to life in an instant. As a result, we have 3D designs that are easy to use, a modeled venue, personalized landing and registration pages, animated graduating avatars, and a support desk.

In this AR/VR-based comprehensive framework for virtual convocation, the main focus is to supply a platform to enjoy the ceremony virtually alongside being user-friendly. So, to achieve the mentioned objectives, it was decided to use Unity3D software for the environment development, VS-code for basic C-sharp coding, Autodesk, 3ds Max for creating and animating 3D models and audition tool for audio purposes. Creating a cross-platform environment is the primary goal of this project. This will help the graduates be engrossed in the ceremony. This project is about creating a real-life experience for the students who are not able to attend their graduation ceremony.

The basic hierarchy includes a ceremony hall designed in Unity3D engine. The students are able to absorb all the attention of the ceremony and enjoy the prestigious moment of life. So creating this virtual convocation ceremony begins with designing a ceremony hall, followed by character design and modeling and providing 3D effects to make it more realistic.

2 Objectives

1. To build a virtual platform (application) for pc/mobile using Unity3D engine along with some coding.
2. To efficiently handle multiple users at a time so that students can enjoy convocation ceremony without any interruption.
3. To make the app as user-friendly as possible for all kinds of users.
4. To try to provide the option to download the e-certificate during the ceremony.

3 Literature Review

In the literature [1], the author discussed the various aspects of virtual reality, its various advantages, and its usage in various engineering sectors. They discussed how virtual reality can help improve the quality of work and provide a better environment for engineers. VR is an immersive environment that uses 3D technology. Virtual reality is a technology in which computer-generated images are used to create a virtual universe. However, it is not a substitute for real-time teaching and learning. It is widely used in education to improve the efficiency and effectiveness of teaching. VR classrooms have become an integral part of education. This paper presents an overall integration solution that combines VR composition with its various features.

In the literature [2], the author talks about virtual reality (VR) technology, which is mainly used in various fields such as entertainment and education. Its various features make it a promising technology for the future. The development of virtual reality technology has many potential effects on users. The paper seeks to clarify the impacts and effects of the technology on users. This paper aims to analyze the various aspects of the VR experience offered by various VR platforms. It also aims to analyze the various factors that influence the VR user experience in various areas. This paper aims to provide suggestions for improving the performance of this technology and to avoid its harmful effects.

In the literature [3], the author has focused on the significant importance and usage of virtual reality in the engineering sector like design, manufacturing, inspection, tooling, assembly, prototyping, etc. In this, virtual reality has been defined as an emerging technology that can provide the user with an actual working environment.

In the literature [4], the author has focused on the three-dimensionality of virtual reality structure. Virtual reality stands apart from other representational technologies because of its unique properties of immersion, presence, and interactivity. Virtual reality does not mimic actual life and does not serve as a representational tool. A human being has the inability to distinguish between perception, hallucination, and illusions. VR has grown into a new phase and becomes a distinct field in the world of computing. The use of virtual reality in automotive design, robot design, medicine, chemistry, biology, education, and building design and construction has already been investigated.

4 Existing System Architecture

This project is inspired from the convocation ceremony held by the IIT Bombay using software's like iclone, Unity3D, etc. Many students used given software to watch the event online and had their e-avatars go up the dais to collect medals and degrees from virtual models of dignitaries and the director. To prevent crowds and minimize the spread of the COVID-19 epidemic, a virtual convocation was organized. With

reference to this idea, this project focuses on providing an e-convocation platform by using Unity3D as the software engine.

5 Proposed Architecture

In this AR/VR-based comprehensive framework for virtual convocation, the main focus is to supply a platform to enjoy the ceremony virtually alongside being user-friendly. So, to achieve the mentioned objectives, it was decided to use Unity3D software for the environment development, VS-code for basic C-sharp coding, Autodesk, 3ds Max for creating and animating 3D models and an audition tool for audio purposes [5]. Creating a cross-platform environment is the primary goal of this project [6]. When the application is launched, the user must register for it and complete the sign-up process. Logging into the app with the proper credentials will be sufficient if the user has previously registered credentials, but if any user has difficulties login in, they will just reset their password by email verification. There are three main steps in creating an animated humanoid character from scratch: modeling, rigging, and skinning.

Modeling: It is the process of creating your own humanoid mesh in a 3D modeling package—3DSMax, Maya, Blender, etc. Although this is a whole subject in its own right, there are a few guidelines you can follow to ensure a model works well with animation in a Unity project.

Rigging: It is the process of creating a skeleton of joints to control the movements of your model.

Skinning: It is the process of attaching the mesh to the skeleton. After character building is completed, it is necessary to add mesh and collider agents that work as barriers so that different objects do not pass through each other.

Then textures and materials are added to every object to make it look realistic. For animation of characters, it is necessary to add unity inbuilt packages such as animator controller, nav-mesh agent, Third person character script, and AI character control script. Using the animator controller, we can create a sequence of different movements (walking, handshakes, etc.), and using imported C# scripts mentioned before, we can control the speed of animation and the location till which it will take place (Fig. 1).

Referring to the code mentioned, if we consider the walking animation, first the character is given a definite path through which it will walk. This is done by placing multiple empty game objects along the path designated, and then, using if else conditions, the character is set to walk from one point to another.

As soon as the users enter the correct login information, they are taken to the homepage. A person's character can be viewed, and they can select the ceremony room that suits their department. They will be redirected, after choosing the venue, to the convocation hall and the ceremony will commence.

In the ceremony hall, 3D models of dignitaries will be seated on the couch and the host will stand near the announcement podium [7]. As the hall will have auditorium

```
using System;
using UnityEngine;

namespace UnityStandardAssets.Characters.ThirdPerson
{
    [RequireComponent(typeof (UnityEngine.AI.NavMeshAgent))]
    [RequireComponent(typeof (ThirdPersonCharacter))]
    public class AICharacterControl : MonoBehaviour
    {
        public UnityEngine.AI.NavMeshAgent agent { get; private set; }             // the navmesh agent required for the path finding
        public ThirdPersonCharacter character { get; private set; } // the character we are controlling
        public Transform target;                                    // target to aim for


        private void Start()
        {
            // get the components on the object we need ( should not be null due to require component so no need to check )
            agent = GetComponentInChildren<UnityEngine.AI.NavMeshAgent>();
            character = GetComponent<ThirdPersonCharacter>();
            agent.updateRotation = false;
            agent.updatePosition = true;
        }


        private void Update()
        {
            //Debug.Log(agent.remainingDistance);
            if (target != null)
                agent.SetDestination(target.position);

            if (agent.remainingDistance > agent.stoppingDistance)
                character.Move(agent.desiredVelocity, false, false);
            else
                character.Move(Vector3.zero, false, false);
        }


        public void SetTarget(Transform target)
        {
            this.target = target;
        }
    }
}
```

Fig. 1 Code snippet for character movement

style seating, the characters of students and special guests will have a seat sequentially when the host calls out the first name, the 3D model assigned to that students will stand up and start moving toward the stage. Once the model walks up the stage and reaches in front of the dignitaries, the character will shake hands, receive the certificate, pause for some time to capture the moment, and exit the stage (just like we observe in real life). Similarly, after every name of the student announced by the host, their 3D model will repeat the same process as mentioned before [8].

A great addition to it can be a system set up inside the application that allows the students to download their certificate's soft copy by displaying a pop-up window on their application screen which will include an identity re-confirmation portal which can be used to monitor the overall attendance. After completing this procedure, the students can then enjoy the remaining convocation ceremony [9]. By using this application many universities which have distance and open learning students, may be able to give a lively ceremony to the students that graduate. Also, as a future improvement of this application is that one can plan to provide environments for different [10] ceremonies required by the users and mentioned in the feedback portal (Figs. 2 and 3).

Fig. 2 Convocation ceremony room

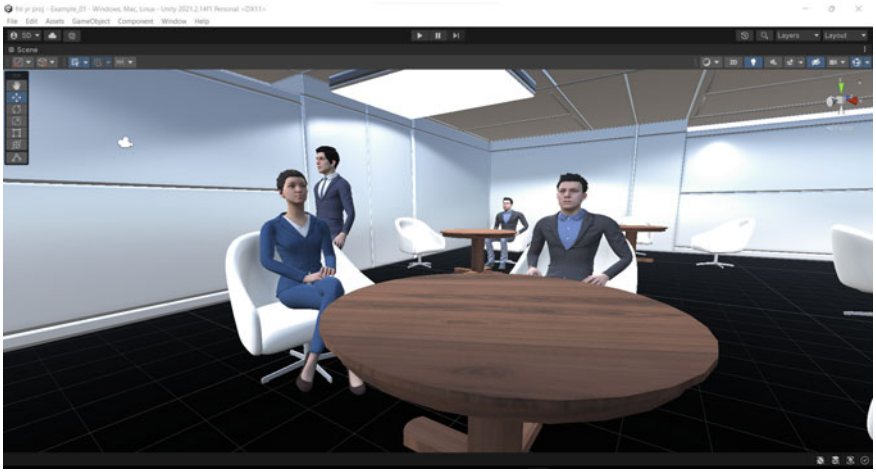


Fig. 3 3D models

6 Conclusion

The process of creating a virtual convocation ceremony and the resources required to achieve it are proposed in this paper. Designing and texturing of the 3D models and objects, along with animation, are the core modules to complete the objectives. The final product of our proposed system will be useful in the future when, due to any uncertainties, people cannot attend any given occasion. Apart from conducting the convocation ceremony, this platform can also be used for family and corporate events. Currently this framework is based on a specific environment with limited user

capacity, considering this limitation, a more developed and efficient system which can handle more characters and animation can be achieved.

The purpose of this venture is to make college students feel the experience of success and satisfaction of graduating virtually. It can also be useful for students who are acquiring distance education. The main objective of this project is to provide a virtual environment using Unity3D and some concepts of virtual and augmented reality referred to in some research papers.

References

1. Mueller K, Bowman D (2020) Introducing the IEEE virtual reality 2020 special issue. IEEE Trans Visualization Comput Graphics 26(5):iv–v. https://doi.org/10.1109/TVCG.2020.2978971
2. Rosenblum L (2000) Virtual and augmented reality 2020. IEEE Comput Graphics Appl 20(1):38–39. https://doi.org/10.1109/38.814551
3. Mueller K, Schmalstieg D (2019) Introducing the IEEE virtual reality 2019 special issue. IEEE Trans Visualization Comput Graphics 25(5):iv–iv. https://doi.org/10.1109/TVCG.2019.2902964
4. Yung R, Khoo-Lattimore C. New realities: a systematic literature review on virtual reality and augmented reality. https://doi.org/10.1080/13683500.2017.1417359
5. Houjeir R, Gotthelf K (2016) The 2016 Dubai ACBSP Region 8 conference at: Dubai, UAE: virtual reality as an education tool
6. Pezeshki Z, Soleimani A, Darabi A (2017) 3Ds MAX to FEM for building thermal distribution: a case study. In: 2017 3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS), pp 110–115. https://doi.org/10.1109/ICSPIS.2017.8311599
7. Shen W, Zeng W (2011) Research of VR modeling technology based on VRML and 3DSMAX. Proc Int Conf Comput Sci Netw Technol 2011:487–490. https://doi.org/10.1109/ICCSNT.2011.6182002
8. Kuang Y, Bai X (2018) The research of virtual reality scene modeling based on Unity3D. In: 2018 13th International Conference on Computer Science and Education (ICCSE), pp 1–3. https://doi.org/10.1109/ICCSE.2018.8468687
9. Wang S, Mao Z, Zeng C, Gong H, Li S, Chen B (2010) A new method of virtual reality based on Unity3D. In: 2010 18th international conference on geoinformatics, pp 1–5. https://doi.org/10.1109/GEOINFORMATICS.2010.5567608
10. Kim SL, Suk HJ, Kang JH, Jung JM, Laine TH, Westlin J (2014) Using Unity3D to facilitate mobile augmented reality game development. In: 2014 IEEE World Forum on Internet of Things (WF-IoT), pp 21–26. https://doi.org/10.1109/WF-IoT.2014.6803110

Oil Spill Detection in Ocean Using Deep Learning

Vighnesh Anand, Aarohi Patni, and Suresh Sankaranarayanan

Abstract Oil spill is a type of pollution which affects both directly and indirectly the human beings, economy of the nation, and the marine life. Oil on top of the ocean damages numerous aquatic organisms since it stops sunlight that is sufficient in achieving the surface of the ocean and lowers dissolved oxygen levels. Oil-coated birds and marine mammals can perish hypothermia because crude oil destroys the insulation. In addition, ingested oil is poisonous to affected creatures and habitat. The answer towards the above-mentioned issues is towards building a methodology employing deep learning towards precisely recognizing oil spills and oil-like spill from ocean for taking appropriate action. So, we in this paper have deployed different pre-trained deep learning models towards classification of oil spill in ocean. In addition, different deep learning models performance are compared and validated in terms of accuracy and losses for proposing the best deep learning model for classification of oil spill in ocean.

Keywords Oil spills · Oil-coated · Image processing · Deep learning

1 Introduction

Oil spills that happen because of mishaps, deliberate dumping, or characteristic reasons, tend to be destructive to marine biological systems. The various reflections demonstrate the simplicity of satellite imagery using synthetic aperture radar (SAR) with respect to the location of the oil spill. Deepwater skyline spill analysed a few coastal SAR images and concluded that most of the oil spill was distinguishable from these images. These sensors can filter large ranges which are not influenced by clouds. They do not rely on sunlight and so it may be recognized where it seems a dark location by the cross-section normalized radar (NRCS). Hydrocarbon compounds decrease the surface roughness of liquid. Dark formation in image data set could be oil spills, organic oil slicks, fronts, land reserves, shear algae, dense ice, inland waves,

V. Anand · A. Patni · S. Sankaranarayanan (✉)
Department of Computer Sciences, College of Computer Science and Information Technology, King Faisal University, Al Ahsa, Al Hofuf, Saudi Arabia
e-mail: pessuresh@hotmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_35

downdrafts and eddies, low wind zones, and rainfall cells. Therefore, one should be exact if the dark formation is oil spill or not. Gulf of Mexico oil spill and Maritime oil spill were certainly caused by an explosion 20 miles off Louisiana and the subsequent sinking so far. This is the largest oil spill so far. This catastrophic situation created the need to develop a highly accurate design that distinguishes between oil spills and oil spill alike.

Satellite images identify oil spills in the ocean in transit. Image processing techniques usually do not classify the severity of oil spill images which can be performed precisely using machine learning or deep learning. There has been work where machine and deep learning techniques have been employed in classification of oil spill in ocean. Most of deep learning models employed are very simple like deep convolutional neural network and VGG19. They have not applied advanced deep learning models to achieve higher accuracy for detection of oil spill in ocean. Oil spills are always addressed with great interest, as small amount of oil can cause harm. We in this research have proposed employing advanced deep learning models in classification of oil spill in ocean. Also, the pre-trained deep learning models employed are validated in terms of accuracy and loss for proposing the best deep learning model for detection of oil spill in ocean. The rest of paper is organized as follows. Section 2 gives the literature review pertaining to algorithms and methodologies used in detection of oil spill using machine learning and deep learning Sect. 3 gives proposed work and methodology towards deep learning models for oil spill detection oil spill in ocean. Section 4 gives the results of proposed work with data set description and metrics towards performance of deep learning models towards oil spill detection. Section 5 gives the conclusion and future work.

2 Literature Review

The research work on oil spill detection in ocean using machine and deep learning has been reviewed in this section.

Authors [1] have developed artificial neural network model towards oil spill detection. The authors here have employed two ANN networks—One for image segmentation and other for feature classification. The data set used for the work was minimal for training. The challenge in this work is that two features gave similar value, and therefore, classification was very difficult. The parameters like wind analysis were not taken into consideration which ultimately affected the accuracy. The parameters taken were statistical parameters which were area, perimeter, and dark spot mean.

Authors [2] in this work developed a methodology for oil spill detection in "synthetic aperture radar (SAR)" images automatically. The algorithm used in methodology here consists of three parts which are dark spot detection, extraction of features from dark spot, and dark spot classification as oil spill or oil spill alike in ocean. The algorithm has been trained on large data set of "Radarsat and Envisat Advanced Synthetic Aperture Radar (ASAR)" images. The performance of the algorithm has been compared with manual and semi-automatic method with benchmark study of

59 “Radarsat” and “Envisat” images. The automatic system performance was faster than manual system with an average processing time of 1.45 min as compared to 10 min for manual inspection.

Researchers in [3] have proposed a boosting approach towards minimizing the misclassification in oil spill. The work here employed data set from 20 SAR images which were separated into 10 classes of oil spill and other 10 classes of oil spill look alike. The models like KNN, two neural models like radial base function and multilayer perceptron, and finally boosting algorithms like AdaBoost and LogitBoost applied. The analysis showed that LogitBoost had better variance compared to AdaBoost.

Authors in [4] have given an overview of various methodologies and techniques used for oil spill detection. “SAR” images were used for oil spill detection. The methodologies used were manual and automatic approaches for oil spill detection. The concentration was given to probabilistic, neural, and fuzzy classification methods. The most common approach for detection of dark formation on SAR images, features extracted from detected dark formation, and classifiers used was discussed.

Authors in [5] have applied clustering, logistic regression, and traditional convolution neural network for detection of oil spill from Envisat and Sentinel-1 satellite images. From the analysis, it was proved that CNN achieved better performance as compared to clustering and logistic regression with lower computation time and mean precision of 0.950.

Authors in [6] have proposed deep convolution neural network called oil spill convolution neural network (OSCNET) for detection of oil spill from “SAR” images. OSCNET is obtained from VGG16 convolution neural network by adjusting hyperparameter with data set of “SAR” dark patches. “OSCNet” consisted of 12 layers which was benefitted from data augmentation for big data set. The performance of “OSCNET” was compared with “AAMLP”—a traditional sophisticated ML classifier. Traditional algorithms are less accurate and cannot handle complex functions and oil spill extraction detection.

Authors in [7] developed deep convolution neural network models towards automating the detection of contaminated areas with associated entities such as twins, ships, or coastal areas. In addition, initial results based on the analysis performed indicate that these models can provide an accurate assessment of an impending disaster with a focus on better situational awareness of the relevant authorities. In this way, these models can be integrated into broader frameworks for disaster and crisis management.

3 Oil Spill Detection Using Deep Learning

Oil spills are dangerous not just to industrial workers, but also to marine. Many authors have employed deep learning and machine learning methodologies for detection of oil spill in ocean and classification of oil spill and oil spill alike in ocean.

There has also been work towards image segmentation towards oil slick into shorter segments for better classification. In terms of seep CNN models, most of the work have focused on VGG16 and nothing beyond that. Also, these deep CNN models are compared with machine learning models like clustering and regression. These models have achieved an accuracy of 94–95% and not beyond that. So, now with the advent of advanced deep learning models, we have validated deep CNN models for better classification of oil spill and oil spill alike in ocean with higher accuracy and reduced error. So, in the forthcoming sections, we give brief idea on simple CNN followed by pre-trained CNN models like VGG19, Inception V3, and Xception used for oil spill classification followed by results and analysis.

3.1 Simple Convolution Neural Network (CNN)

CNN is a neural network with several convolutional layers (and several other layers). The convolution layer has a set of filters that perform convolution operations. CNN is a class of artificial neural networks most used for visual image analysis. "CNN" consists of an "input layer, a hidden layer, and an output layer". CNN takes raw image pixel, trains the model, and automatically extracts the features for better classification.

3.2 VGG19

"VGG" is the successor to "AlexNet", but it was named "VGG" because it was created by another group called the "Visual Geometry Group" in Oxford. It carries and uses some ideas from its predecessor and uses deep convolutional neural layers to improve them and improve accuracy. The "VGG19" is a variant of the "VGG" model, consisting essentially of 19 layers (16 convolution layers, 3 fully connected layers, 5 MaxPool layers, and 1 SoftMax layer). There are other VGG variations such as VGG11 and VGG16.

3.3 Inception V3

"Inception V3" is a commonly used image recognition model which achieved higher performance of 78.1% accuracy. The model results of different ideas explored by researchers over time. The model consisted of "convolutions, average pooling, max pooling, concatenations, dropouts, and fully linked layers" which are building blocks of symmetric and asymmetric components. Batch normalization is applied towards activating the inputs and used extensively throughout the model.

3.4 Xception

"Xception" is a 71 layer deep convolutional neural network which imports a pre-trained version of network from "ImageNet" database. This has been trained over a million images. This model is capable of classifying 1000 different image objects. "Xception" is an expansion of the "Inception" architecture that uses depth-wise separable convolution to replace the regular Inception modules.

3.5 DensNet-169

DenseNet-169 model is a member of the DenseNet family of image classification models. The DenseNet-169 is slightly larger, weighing at around 55 MB vs the DenseNet-121's 31 MB. "Convolution neural network" with smaller connection between layers which are adjacent to input and closer to output layer can result in deeper, accurate, and more efficient in training. "DenseNet" which is a "feed forward" neural network connects every layer to other. In this model, "feature maps" of the previous layer are used as input to subsequent layers. "DenseNet" got lot of advantages which include eliminating vanishing gradient, propagating the features, reusing the features, and reducing the number of parameters.

3.6 Inception-ResNet v2

"Inception-ResNet-v2" model is trained over a million photographs from "ImageNet" database. The 164 layer network got capability of identifying photos in different object categories. This model is built based on "Inception" model and residual connection. In addition, multiple size "convolution" filters are integrated with the residual connection in the "Inception-ResNet" block. "Inception-ResNet v2" is variant of "Inception V3". These inception blocks have been modified with few parallel towers compared to previous "Inception V3".

3.7 CNN-LSTM

The "CNN long short-term memory network (CNN-LSTM)" starts with CNN layers added for feature extraction followed by "LSTM" layers with a dense layer on output resulting in "CNN-LTM". "CNN-LSTM" can be perceived as two sub-models where "CNN" used for extraction of features from images followed by "LSTM" for interpretation of features of images over time steps. "CNN" would process only one image which converts the input pixels into an internal matric or represented as vector. This

methodology is repeated over different images for "LSTM" to build internal state and updating the weights using "back propagation through time (BPTT)" across a sequence of internal vector representation of input images.

4 Implementation of Deep CNN Models

The implementation of CNN models for classification of oil spill in the ocean requires good number of data set. The data set needed for classification of oil spill in the ocean was not available from any source directly like Kaggle, UCI. So, we need to do web scraping to capture the data set from web for oil spill and non-oil spill from ocean which amounted to 2000 samples. These 2000 samples consisted of 1000 oil spill and 1000 non-oil spill images. So, towards increasing the data set, data augmentation was applied. Data augmentation is a technique towards increasing the number of data sets by applying techniques like flipping, rotating, cropping, scaling and so forth. This results in slightly modified copies of already existing data set or creating new synthetic data from existing data set. This resulted in total of 4000 samples of data set consisting of 2000 oil spill and 2000 non-oil spill. So, from the augmented data set, 2800 samples consisting of 1400 of oil and non-soil spill taken for training. Remaining 600 samples which consisted of 300 oil spill and 300 non-oil spill taken for validation. The rest of 600 samples which consisted of 300 oil spill and 300 non-oil spills taken for testing. That is 70% taken for training and remaining 30% taken for validation and testing.

5 Performance Analysis of Deep CNN Models

For deep learning model, all images were resized to a dimension of 224 * 224 * 3 for VGG19 and 150 * 150 * 3 for CNN, CNN-LSTM, Inception V3, Xception model, DenseNet-169, and Inception-ResNet v2 as input shape Further to that, the pre-trained models like CNN, VGG19, Inception V3, and Xception were applied which comprised of hidden layers and an output dense layer with SoftMax activation. The optimiser used for all models is Adam. The performance of all models in terms of accuracy, precision, recall, F1 score, and losses is tabulated (Tables 1, 2, 3, 4, 5, 6, 7, and 8).

Table 1 CNN model performance

Model		Precision	Recall	f1-score	Accuracy
Simple CNN	Non-oil spill (Class 0)	0.98	1.00	0.99	0.99
	Oil spill (Class 1)	1.00	0.98	0.99	

Table 2 VGG19 model performance

Model		Precision	Recall	f1-score	Accuracy
VGG-19	Non-oil spill (Class 0)	0.99	1.00	1.00	0.99
	Oil spill (Class 1)	1.00	0.99	0.99	

Table 3 Inception V3 model performance

Model		Precision	Recall	f1-score	Accuracy
Inception V3	Non-oil spill (Class 0)	0.99	0.99	0.99	0.99
	Oil spill (Class 1)	0.99	0.99	0.99	

Table 4 Xception model performance

Model		Precision	Recall	f1-score	Accuracy
Xception	Non-oil spill (Class 0)	0.99	1.00	1.00	0.99
	Oil spill (Class 1)	1.00	0.99	0.99	

Table 5 DenseNet-169 model performance

Model		Precision	Recall	f1-score	Accuracy
Densenet-169	Non-oil spill (Class 0)	0.99	1.00	0.99	0.99
	Oil spill (Class 1)	1.00	0.99	0.99	

Table 6 Inception-ResNet v2 model performance

Model		Precision	Recall	f1-score	Accuracy
Inception-ResNet v2	Non-oil spill (Class 0)	0.99	1.00	0.99	0.99
	Oil spill (Class 1)	1.00	0.99	0.99	

Table 7 CNN-LSTM model performance

Model		Precision	Recall	f1-score	Accuracy
CNN-LSTM	Non-oil spill (Class 0)	1.00	1.00	1.00	1.00
	Oil spill (Class 1)	1.00	1.00	1.00	

Table 8 Model loss results

Models	Loss
CNN	0.09
VGG19	0.01
Inception V3	0.1
Xception	0.06
DensNet-169	0.04
Inception-ResNet v2	0.01
CNN-LSTM	0.08

From the analysis of model, it is found that all the deep learning models have achieved an accuracy of 99% towards classification of oil spill and non-oil spill class. In terms of precision, recall, and F1 score, it has been found that 98% and above which signifies that number of false positive and false negative been less resulting in effective classification. Now in terms of loss, VGG19 followed by Inception-ResNet v2 model resulted in least loss during validation phase for the number of epochs as compared to other models.

6 Conclusion and Future Work

So, we in this paper have compared different deep learning models like CNN, VGG19, Inception V3, and Xception on oil spill and non-oil image data set resulting in higher accuracy and minimal losses. The data set captured consisted was augmented by applying different techniques resulting in 4000 images consisting of 2000 oil spill and 2000 non-oil spill. Though other literature has applied machine learning and deep learning models, their accuracy achieved was lesser compared our models. Also, the models used were not advanced deep learning models like Inception and Xception models. In future, we would be integrating the deep learning models for achieving higher accuracy and applying explainable AI like LIME for explanation of model classified.

References

1. Singha S, Bellerby TJ, Trieschmann O (2013) Satellite oil spill detection using artificial neural networks. IEEE J Selected Topics Appl Earth Observations Remote Sens 6(6):2355–2363
2. Solberg AHS, Brekke C, Husoy PO (2007) Oil spill detection in radarsat and envisat SAR images. IEEE Trans Geosci Remote Sens 45(3):746–755
3. Ramalaho GLB, Medeiros FNS (2006) Using boosting to improve oil spill detection in SAR images. In: Proceedings of 18th international conference on pattern recognition. IEEE, Hong Kong, China, pp 1–4
4. Topouzelis KN (2008) Oil spill detection by SAR images: dark formation detection, feature extraction and classification algorithms. Sensors 8(10):6642–6659
5. Cantorna D, Dafonte C, Iglesias A, Arcay B (2019) Oil spill segmentation in SAR images using convolution neural network . a comparative analysis with clustering and logistic regression algorithms. Appl Soft Comput 84:1–13
6. Zeng K, Wang Y (2020) A deep convolutional neural network for oil spill detection from spaceborne SAR images. Remote Sensing 12(6):1–123
7. Krestenitis M, Orfanidis G, Ioannidis K (2019) Early detection of oil spills in satellite images using deep CNN. In: Proceedings of 25th international conference on multimedia modeling. Springer, Thessaloniki, Greece, pp 1–12

Roof Garden Irrigation and Drainage Automation Using Microcontroller

Md. Shiam Prodhan, Nazmuj Shakib Diip, Mir Nushrat Zahan Tirumony, Md. Afraim Bin Zahangir, and Bijan Paul

Abstract The roof garden irrigation and drainage system is a take on the traditional roof gardening to ease the strenuous effort and help the environmental balance. As we know while more and more buildings are getting constructed, people are losing valuable places for gardening so to overcome this, people are now doing roof gardening which is a hobby that is also good for the environment and to make people more encouraged to do roof gardening, we proposed this system. The task has been accomplished using an Arduino Uno (ATmega328P), a soil moisture sensor communicating the primary moisture status of the soil, two water pumps for irrigation and drainage each, and NodeMCU for wireless control. This device tends to help in the daily irrigation of multiple tops/surfaces at the individual required time and drain excess water which is harmful to the plant and environment if kept unattended. This device also opens a platform for further studies on, for example, the water that is drained can be used to see the number of harmful products that products contain and how this water can be harmful to surroundings and plants.

Keywords Automatic irrigation · Smart drainage · Wireless control

Md. S. Prodhan · N. S. Diip · M. N. Z. Tirumony · Md. A. B. Zahangir · B. Paul (✉)
University of Liberal Arts Bangladesh, Dhaka, Bangladesh
e-mail: bijan.paul@ulab.edu.bd

Md. S. Prodhan
e-mail: shiam.prodhan.cse@ulab.edu.bd

N. S. Diip
e-mail: nazmuj.shakib.cse@ulab.edu.bd

M. N. Z. Tirumony
e-mail: nushrat.zahan.cse@ulab.edu.bd

Md. A. B. Zahangir
e-mail: afraim.zahangir.cse@ulab.edu.bd

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_36

1 Introduction

The practice of roof gardening was started in 1463 [1] and with that started many new fields and concepts. With the advancement of time and technology, the culture of roof gardening has faced many challenges and caused some problems as well. Some of the challenges being, it takes a lot of effort and time. Over-irrigation is not only wasting water but also is deadly for plants, stagnation of water becomes home for pests are some of the problems of roof gardening in today's world. Also, people like to do roof gardening. They like to water their plants in the morning in the beginning perhaps but as time passes it becomes a problem for them as these days it is very difficult to make free time and attend to your plants. The solution to these is found using computational automation and logic.

This Arduino-based automation system (Figs. 1 and 2) discusses an automatic irrigation system based on the moisture content of the soil using the soil moisture sensor placed inside the pitcher. The flow of water is controlled by the water pump motor, which is controlled by the output of the sensors. Similarly, a secondary sensor is placed on the top edge of the pitcher which senses the overflow of water, that is, excess water, and triggers a different pump to be activated and suck out the excess water from the pitcher to a safe reserve container. The drained water is not thrown away rather is kept in a safe container for later use.

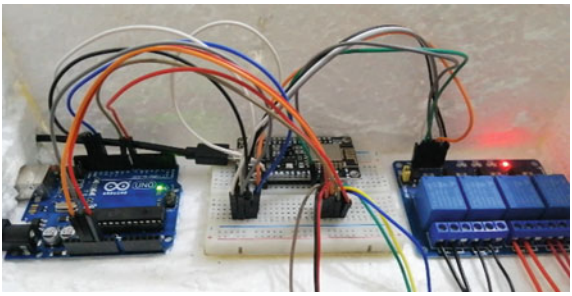


Fig. 1 Main circuit of the system

Fig. 2 Full device at work

It is discussed here the accomplishing areas of, the automatic control of the water flow based on the moisture level of the soil, understanding the mechanics, automation algorithms, plant-based research fields, and possibilities of IoT in the field of agriculture.

This paper discusses the related works in Sect. 2. Section 3 holds the record of accomplishments achieved by this research work. Section 4 talks about the system architecture and Sect. 5 talks about the system design of the research work. The programming is discussed in Sect. 6. The future possibilities are focused on Sect. 7. The conclusion was drawn in Sect. 9.

2 Related Works

Vassureshvs1995 has presented a project about IoT-based automation and monitoring system on a website called my device cayenne community. The system is prepared to sense the amount of light falling on the plants, soil moisture, and water flow rate. When the moisture sensor can sense that soil is too dry, then the system will give the command to start a pump for watering soil. The flow meter monitors the water consumption. In this system, you can monitor all the information from your smartphone [2].

Pareena Jariyayothin, Kachaporn Jeravong-aram, Nattakarn Ratanachaijaroen, Thitinan Tantidham, and Puwadech Intakot have done another project which was same about smart monitoring system and controlling over homegrown plants. They used a mobile application to know the update of plant growth. The main advantage of that system was the smart monitoring system with the application and implementation and control of the watering process for plants that are growing. They used NodeMCU V3 from which the used to send and receive data and transfer the data into a microcontroller to monitor the watering sensor perfectly [3].

T. Thamaraimanalan, S. P. Vivekk, G. Satheeshkumar, and P. Saravanan have presented the smart garden monitoring system where they used IoT in their paper and published it in the Asian Journal of Applied Science and Technology (AJAST) where the final object of that project was to maintain the nature of growing plants by regularly monitoring the criteria that leads to the increasing life of both plants system and human beings. This automatic system works as a manual system. By using the android system, they created mobile applications that used to monitor all the parameters of plants in each time of the garden and checkout automatically the watering procedure. Also, NodeMCU has been used for collecting the parameters of the soil and transfer the information into a firebase through an inbuilt Wi-Fi system [4].

By building IoT plant monitoring system with smart technology the Team Bolt-Makers: Pranav Pai Verneker, Devashis Gupta can monitor water content and sunlight on trees and send alerts when its actions. But they did not use any coding just drag and drop the logic. They have used soil moisture sensors and some software apps and online services for massaging [5].

Again an IoT-based garden monitoring system that was presented by M. Sambath, M. Prasant, N. Bhargav Raghava, S. Jagadeesh that is published in a conference paper on Physics and Photonics Processes in NanoSciences is designed by moisture sensors for suitable moisture levels, and an automatic water system has been developed a less quantity of water consumption [6].

Tanu Sahu and Ashok Verma have been prepared an automated monitoring and watering system that will work for water scarcity and reduce the excess water also. The developed system with an automatic sprinkler will work for distributing water to all plants and never waste any water. This system will help to collect data, such as temperature, water needs by using a soil moisture sensor, Raspberry Pi, and will be notified to the user [7].

Ashphiar Raihan has published a project idea about smart gardening on the project hub website on November 20, 2016. In that project, by using soil moisture it can scene the water level, temperature, and humidity sensor, it can tell the temperature. They also used photo resister for light conditions [8].

R. Nageswara Rao and B. Sridhar have published a paper on the IoT-based smart technological crop-field monitor and automated irrigation system on IEEE Xplore. It is a Raspberry Pi-based automatic irrigation system that works for improving the productivity of crop and crop development at less quantity of water consumption system. This system starts watering the land when needed. The sensors work for sending the data of dry soil to the user, and the system starts watering till the soil gets wet [9].

Dr. Willem Van Cotthem published his project on Conference: First International Summit for Afforestation Roof Gardens in China—April 2005 [10].

Shrinidhi Rajagopal et al. have published an article about the automatic watering system for the plants at a specific time. Based on the Internet of things, the object-oriented design is described in the system [11].

Smart Garden by Patel Dipen is a project about how to maintain a garden in modern life. He uses Arduino and Raspberry Pi, soil sensor, and Blynk app used to view those sensor data from any location [12].

3 Application

This research work shows successful irrigation and drainage of excess water smartly using microcontrollers and sensors. With the help of this system, the problem of over-irrigation is solved which saves the plant from harm. This system gives a successful result on smart drainage which solves the problem of stagnant water. A simple solution for water wastage due to human error during manual irrigation is met smartly using machine-controlled automation in the irrigation system.

This research work also creates the possibility of further study in the field of biology and agriculture. The daily water consumption of different plants is based on geolocation, plant types, altitudes, and temperature.

4 System Architecture

Once the Arduino is powered up, the sensor is enabled and it sends the information about the water level at the top. Depending on the information, Arduino decides whether to remove the excess water or transfer water to the pot. Also, by using the NodeMCU, the device can be controlled by using a mobile app (Fig. 3).

- For the watering system, when no water will be present in flower pots, then the motor will be turned on and fill the pot with water up to a certain level, then it will stop. These can also be done using the NodeMCU.
- For the drainage system, when excess water will be present in flower pots, then the motor will be turned on and remove excess water from the pots after that it will stop. These water drainage processes can also be done using the NodeMCU.

Device	Use
Arduino	To implement the required codes
Water pump	To transfer the water
Relay switch	Breaking the circuit
Adaptor	To turn AC into DC voltage
Breadboard	For making parallel connections
Jumper wire	For making connections between devices
External power source	To power up the Arduino
NodeMCU	To control the device via an app

5 System Design

The circuit diagram shows the practical flow of the flowchart diagram and gives a visual representation. The NodeMCU is controlled via Wi-Fi with the Blynk app, a free open source IoT device controlling app for both Android and iOS devices (Fig. 4).

- For the watering system: Sensor 1 reads the data and sends it to Arduino. If the voltage is high, then Arduino turns off the relay switch which breaks the circuit, and the motor 1 is turned off. If the voltage is low, then the relay switch is turned on, and the motor is also turned on. Similarly, NodeMCU can also be used to do the same functions by voice or tapping a button.
- For the drainage system: Sensor 2 reads the data and sends it to Arduino. If the voltage is high, then Arduino turns off the relay switch which breaks the circuit, and the motor 2 is turned off, and if the voltage is low, then the relay switch is turned on, and the motor is also turned on. Similarly, NodeMCU can also be used to do the same function by voice or tapping a button (Figs. 5 and 6).

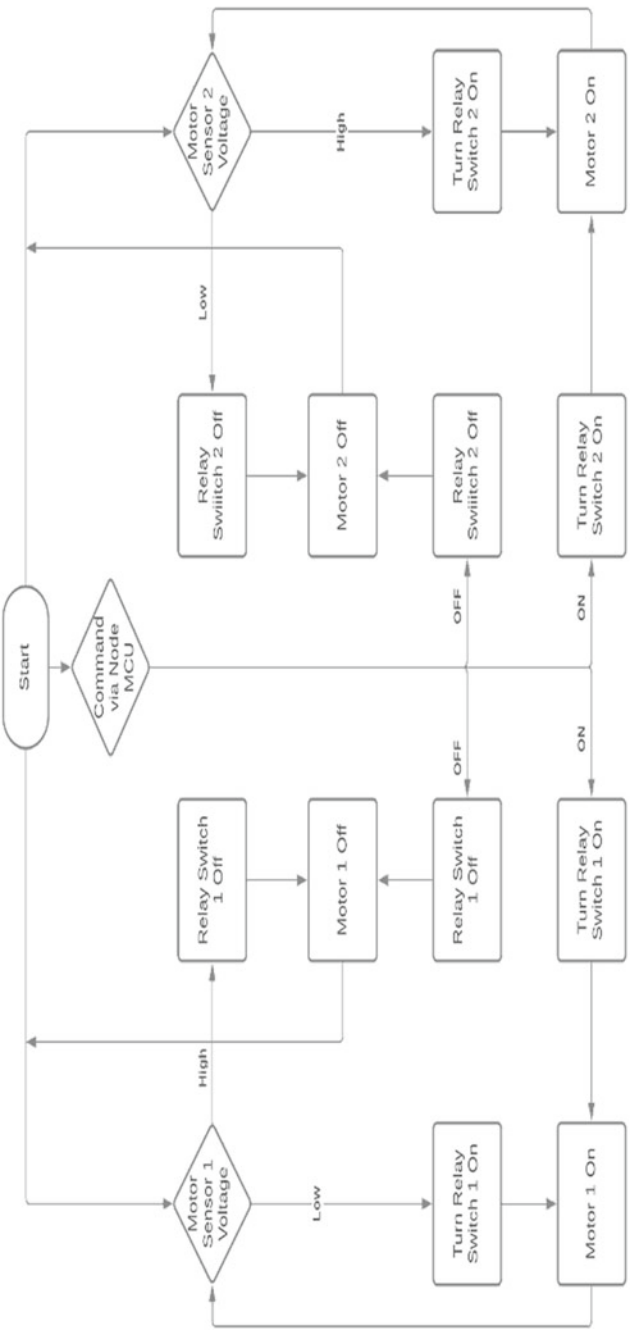


Fig. 3 Architecture of the system

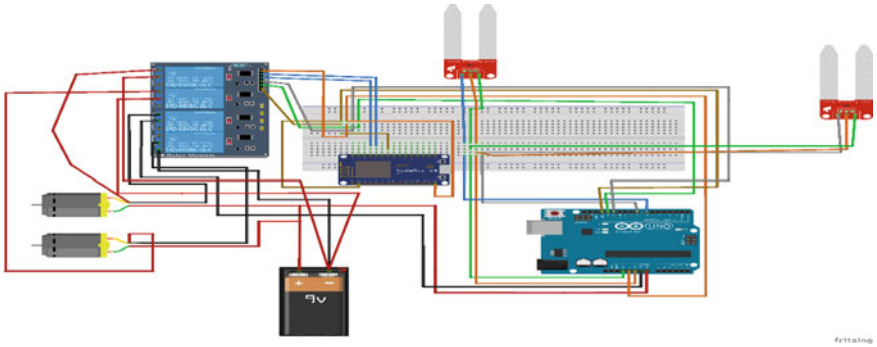


Fig. 4 Circuit diagram of the system

Fig. 5 Controlling using app

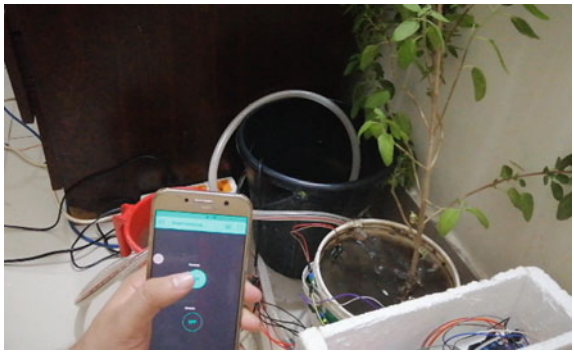
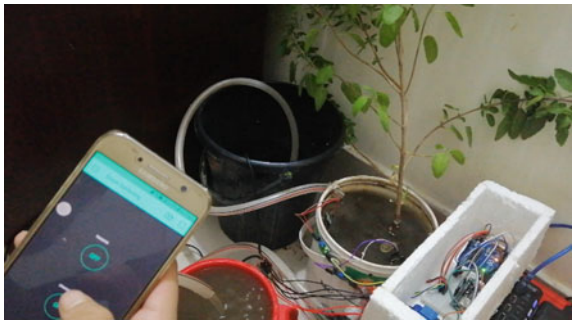


Fig. 6 Controlling using app



6 Programming

The pseudocode of the system is discussed below:

Program Starts
Initialize variable ACWATERPUMP1 = 13
Initialize variable sensor1 = 8

Initialize variable ACWATERPUMP2 = 12
Initialize variable sensor2 = 7
Initialize variable val
Start setup function Call function pinMode with input 13 and OUTPUT
Call function pinMode with input 8 and INPUT
Call function pinMode with input 12 and OUTPUT
Call function pinMode with input 7 and INPUT
End setup function
Start loop function Initialize variable val by calling function DigitalRead with input 8
IF val = LOW
Call function DigitalWrite with input 13 and HIGH
ELSE
Call function DigitalWrite with input 13 and LOW
Initialize variable val by calling function DigitalRead with input 7
IF val = LOW
Call function DigitalWrite with input 12 and HIGH else Call function DigitalWrite with input 12 and LOW
Call delay function with 1000ms End loop function
Program Ends

7 Future Works

This research work has huge possibilities for future work. Conversion from microcontrollers to microprocessors can give more output. The use of Raspberry Pi with a cloud server could be done for real-time communication. Additional pH sensors and solar sensors can be used for the collection of data and power. Smart fertilization can be done using AI and machine learning.

8 User Feedback

Organization	Attendance
Dhaka University	15
University of Liberal Arts Bangladesh	30

9 Conclusion

In this paper, we proposed a method for automating the process of roof gardening using a microcontroller, sensors, and actuators to automate the process of water flow. The roof garden irrigation and drainage automation system connects the traditional roof gardening culture with modern technology to ease human effort and time. Additionally, the research work supports in prevention of water wastage, over-irrigation, and stagnation of water. Using the Arduino, sensors, and NodeMCU, the device controls irrigation and drainage of water at the required time automatically. The Blynk app allows the control of NodeMCU wirelessly from smartphone devices. The proposed method is limited to one plant or pot at a time which is not economically optimized at this moment. Also, the water management of the drainage system and the dirt and fertilizer collection and management is not included in this paper. With further research and resources, these limitations can be overcome and an optimized, one-to-many system can be made for the future of smart agriculture.

References

1. Helen S (2017) The history of rooftop gardens. [Online] turfmagazine.com. Available at: https://www.turfmagazine.com/landscape-designbuildinstallation/rooftop-escapes/#:~:text=This%20living%20mountain%20in%20the,York%20City%20in%20the%201890s
2. Vassureshvs1995 (2017) IoT based garden automation and monitoring system. [Online] community.mydevices.com. Available at: https://community.mydevices.com/t/iot-based-garden-automation-and-monitoring-system/2735
3. Pareena J, Kachaporn J, Nattakarn R, Thitinan T, Puwadech I (2018) IoT backyard: smart watering control system. [Online] ieeexplore.ieee.org. Available at: https://ieeexplore.ieee.org/document/8523856
4. Thamaraimanalan T, Vivekk SP, Satheeshkumar G, Saravanan P (2018) Smart garden monitoring system using IOT. Asian J Appl Sci Technol (AJAST) 2(2):186–192 (April–June 2018)
5. Pranav PV, Devashis G (15 May 2018) Build a smart garden with IoT plant monitoring system. [Online] hackster.io. Available at: https://www.hackster.io/boltmakers/build-a-smart-garden-with-iot-plant-monitoring system-3edc09
6. Sambath M, Prasant M, Bhargav Raghava N, Jagadeesh S (2017) IoT based garden monitoring system. In: International conference on physics and photonics processes in nano sciences
7. Tanu S, Ashok V (2017) [Online] ijcaonline.org. Available at: https://www.ijcaonline.org/archives/volume172/number6/28254-2017915160
8. Ashphiar R (20 Nov 2016) Smart garden [Online] project hub. Available at: https://create.arduino.cc/projecthub/smart-boys/smart-garden-2dd7b0
9. Nageswara Rao R, Sridhar B (2018) IoT based smart crop-field monitoring and automation irrigation system. [Online] ieeexplore.ieee.org. Available at: https://ieeexplore.ieee.org/document/8399118
10. Cotthem WV (2015) Rooftop gardening. In: Conference: first international summit for afforestation roof gardens
11. Shrinidhi R, Krishnamurthy V (2017) OO design for an IoT based automated plant watering system. [Online] ieeexplore.ieee.org. Available at: https://ieeexplore.ieee.org/document/7944101

12. Patel D (21 May 2017) Smart garden is a plant environmental system. [Online] PROJECT HUB. Available at: https://create.arduino.cc/projecthub/patelDeep/smart-garden-cda82f?ref=similar&ref_id=27258&offset=0

Smart System to Reduce High-Beam Glare

Al Imran Fakir, Aishwariya Farahi, Jannatul Ferdouse Sornali, Rakib Ahmed, and Bijan Paul

Abstract The highest rate of accidents occurs on the highway roads during the night. Most of the drivers use high beams while driving at night which causes blurred vision of the person coming from the opposite direction. This intense headlight beam hits the driver's eye causing him to lose perception and ability to see objects or movements that are not directly in line with their eyes. To avoid such accidents due to temporary blindness of the drivers, we can use a smart system to reduce high-beam glare. The system will detect the high-beam glare using the sensor and will automatically reduce the glare so that the driver does not get distracted and maintains a clear eye vision while driving at night. This system is important because the accident rate is much higher at night only because of the high-beam glare of vehicles from the opposite direction.

Keywords Accident · High-beam light · Glare · Blurred vision · Peripheral vision

1 Introduction

Light is essential to see. At night, when there is no light, we lose our perception and ability to see objects that are not in line with our vision. Driving at night is one of the most common causes of road accidents. Even though we can see at night,

A. I. Fakir · A. Farahi · J. F. Sornali · R. Ahmed · B. Paul (✉)
Department of Computer Science and Engineering, University of Liberal Arts Bangladesh, Dhaka, Bangladesh
e-mail: bijan.paul@ulab.edu.bd

A. I. Fakir
e-mail: imran.fakir.cse@ulab.edu.bd

A. Farahi
e-mail: aishwariya.farahi.cse@ulab.edu.bd

J. F. Sornali
e-mail: jannatul.ferdouse.cse@ulab.edu.bd

R. Ahmed
e-mail: rakib.ahmed.cse@ulab.edu.bd

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_37

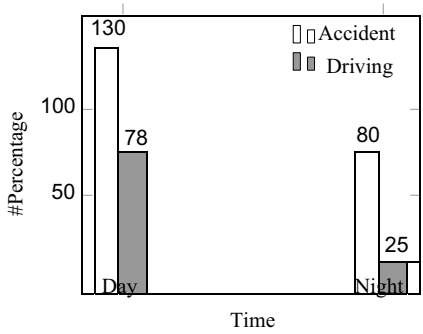


Fig. 1 Statistics of driving versus accident

street lights, headlights, taillight, and darkness mixed together make it more difficult to see. It is hard to adjust our vision when we move from darkness to bright light and bright light to darkness again and again. And this happens a lot when we are looking straight into the light from oncoming vehicles. About 1.3 million people die every year due to road accidents [1]. While only 25% of the vehicles are driven at night, 55% of road accidents happen at night [2]. High beams can be distracting for both drivers and pedestrians. If you ever had to drive at night, you know how hard it is to keep changing between high beam and low beam manually, while neither of them fits the situation perfectly. To prevent these accidents, this system will do it automatically for you and even change the direction of light coming from your headlight when needed. It has a light-dependent resistor (LDR) sensor for detecting the presence or the absence of light. If a place has enough light, the headlight will switch from high beam to low beam. If a place is dark, the headlight will remain in a high beam. There is another light-dependent resistor (LDR) for sensing the high beam of the opposite vehicle. If the vehicle has a high beam, the headlight intensity of our vehicle will reduce so that other drivers can see where they are going (Fig. 1).

IR sensors will detect objects within 4 m proximity, if the high beam is on, then it will rotate downwards with the help of servo motors so that pedestrians or animals can stay out of the way. This technique completely avoids both the drivers from glaring effect if both vehicles have this system installed. If two vehicles 'A' and 'B' are traveling toward each other during the night with their high beam switched on, when both vehicles come close to each other, vehicle A will automatically decrease the intensity of the high-beam light and help the vehicle B driver to see properly. Similarly, if vehicle 'B' has the same system, there is no chance of complete loss of sight and it will save both the vehicle from an accident.

2 Literature Review

Adaptive driving beam headlights (ADB) use beams as bright as high beams. However, instead of changing from high beam to low beam or vice versa, it can

change the shape, brightness, and direction. A shutter system, which covers sections of the headlight beam, is made up of many light sources that can be turned on and off as needed; and can be used to do this [3].

During the lockdown, some students from Kerala introduced a device that can automatically dim the light of their vehicle when light from the opposite vehicle catches the sensor. The device detects other vehicles using RF transceivers, checks whether or not a high beam is switched on using photodiodes, and if a high beam is on, the device changes it to a low beam [4].

An automotive system can also detect other traffic and switch the high beams accordingly by using an image processing system [5].

A prototype of an automatic headlight intensity control system using phototransistors, LCD, IR sensor, and ARM controller [6].

A camera is used to detect oncoming and preceding vehicles to identify their head or tail lights at night time [7].

Arduino Uno is used automatically dim the headlight when the vehicle is approaching in the opposite direction [8].

Light-dependent resistor (LDR) module is used to detect the opposing vehicle by using an Arduino Uno and an ultrasonic sensor [9].

A single hardware architecture programmed to carry out a wide range of functions such as anti-glare headlight beams, enhanced visibility, increased clarity of lanes, signs, and walkways, and early warning of obstacles [1].

In this paper, we developed a simple system that can detect objects within 4 m of proximity using IR sensors, and rotate the direction of light using a servo motor. Light-dependent resistors, LDR 1 and LDR 2, are used to detect light and darkness and help the device automatically switch beams accordingly.

3 Process Diagram

It has a light-dependent resistor (LDR) sensor for detecting the presence or absence of light. If a place has enough light, headlight will switch from high beam to low beam. If a place is dark, headlight will remain in high beam. There is another light-dependent resistor (LDR) for sensing the high beam of the opposite vehicle then it will switch from high beam to low beam so that other drivers can see where they are going. IR sensors will detect objects within 4 m proximity, if the high beam is on then it will rotate downwards with the help of servo motors (Fig. 2).

4 Technical Specifications

See Table 1.

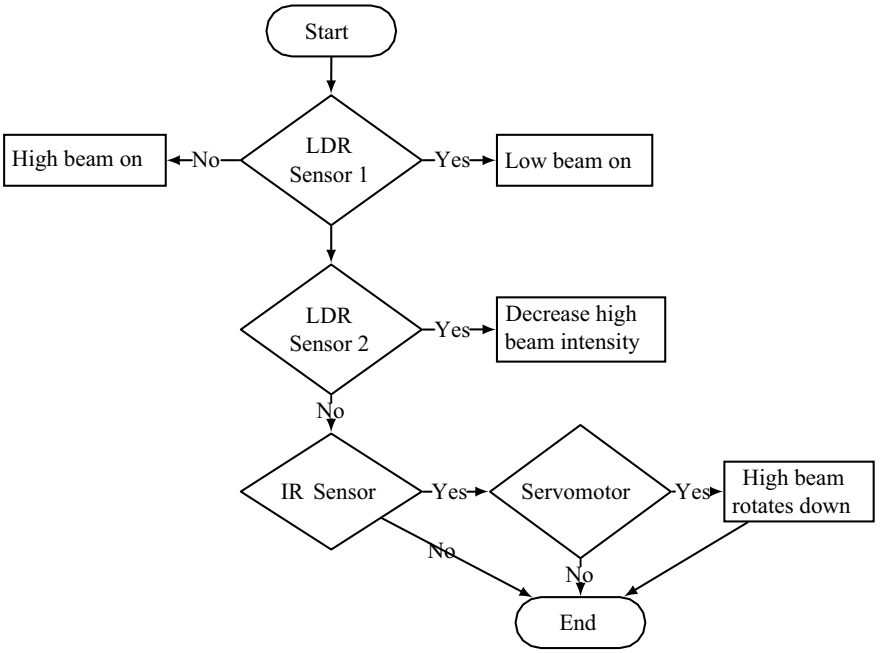


Fig. 2 Flowchart of the system

Table 1 Table of components

Components	Usage
Arduino Uno	For connecting to the computer system
Breadboard	To connect components of the circuits
Light-dependent resistors	To sense the presence and the absence of light
IR sensor	To detect objects and motion
Servo motor	To change directions of the light
Lights	To visualize
Wires	For connection
1 k resistor	To limit current flow

5 Design and Implementation

See Fig. 3.

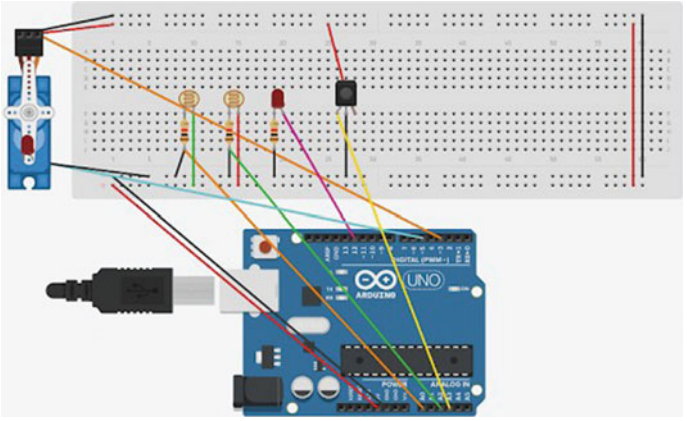


Fig. 3 Circuit diagram

5.1 Circuit Implementation Methodology

LDR 1 sensor is connected to the positive terminal. Another pin is connected to a 1k resistor and the negative terminal of the circuit and the A0 pin of Arduino Uno. LDR 2 sensor is connected to the positive terminal. Another pin is connected to a 1k resistor and the negative terminal of the circuit and the A2 pin of Arduino Uno. Low-beam light's negative terminal is added to the 12th pin of Arduino Uno, and the positive terminal is connected to the negative terminal with a 1k resistor. Servo motor has 3 portions. One is connected to the positive terminal, the other is connected to the negative terminal, and the last one is connected to the 3rd pin of Arduino Uno. The negative terminal of high-beam light connected to the servo motor is added to the 5th pin of Arduino Uno, and the positive terminal is connected to the negative terminal. The IR sensor has 3 portions. The power pin is connected to the positive. The ground pin is connected to the negative. The output pin is connected to the A3 pin.

5.2 Working Proof and Result

From Figs. 4 and 5, it can be seen that when it is enough to light available, then the low beam is activated (Fig. 4). And when there is not enough light available, then the high beam is activated (Fig. 5). Therefore, it can be stated that, the system works perfectly and can make the decision based on the light availability to reduce the high-beam glare.

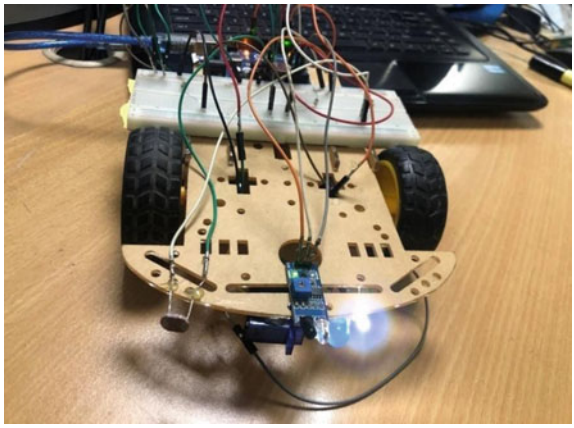


Fig. 4 Low beam in presence of sufficient light

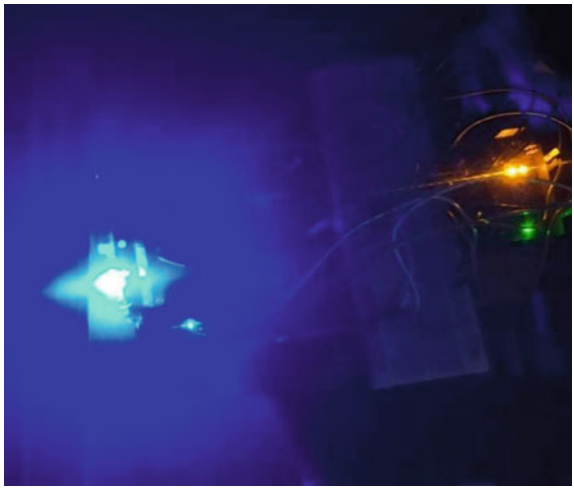


Fig. 5 High beam in absence of light

5.3 Pseudocode

```
Require: IR Sensor readings
Ensure: Proper controlling of the high beam glare
 while vehicle is active do
  irVal = analogRead(irs);
   if 50 < irVal < 400 then
     Go to servomotor;
     if position > 90 then
        Change position because object detected else
        Position won't change because object isn't detected
      end if
   end if
   if IDR2Value > 100 then if
     irValue < 25 then
        Shift to low beam because darkness detected end
      if
   end if
   if IDR2Value < 25 then if
     irValue > 100 then
        Shift to high beam because light detected
     end if
    end if
 end while
```

6 Limitations and Future Goals

This automated system is great, but there are some limitations. For instance, while driving in a residential area, the system may confuse street lights with approaching vehicles, causing it to switch between the two beams when it should not. Another limitation is, the car manufacturers and lawmakers have to maintain guidelines so that the device is used in all vehicles. To overcome these limitations and reduce glare, we can use polarizing filters. Horizontal polarization occurs when light reflects off horizontal surfaces such as the road, water, or snow. A vertically oriented polarizing filter will reduce glare while still allowing plenty of light to pass through. A shutter system is another great alternative, which covers sections of the headlight beam, made up of many light sources that can be turned on and off as needed.

The project's future goals include using this technology in real-time vehicles and building a system that can reduce glare without needing to communicate with other vehicles, improving the system's efficiency with the support of industry professionals, and making it less expensive.

7 Conclusion

As per the records of the past years, vehicular accidents are mostly based on the use of the wrong beam. Through this proposed system, drivers will not feel temporary blindness due to high-intensity headlight glare from the vehicles on the opposite side if both vehicles have this system. It can measure the distance of another vehicle and dim the high-beam light to a low beam when necessary. Using this system will decrease the risk of accidents of drivers and passengers during nighttime. Driving a manual transmission vehicle is riskier than an automatic vehicle because human error causes 94% of all vehicle accidents [10]. If manual or auto transmission vehicle has this system, they can also able to take the facilities of luxury vehicles. By using this system, driver can have a comfortable vision in all conditions during nighttime which features enjoyed in luxury vehicles.

References

1. Road Traffic Injuries, WHO (2021)
2. Night driving, Pines Salomon Injury Lawyers, APC (2021)
3. Chris Teague (2021) Here's how smart headlights work
4. Nagadeesh N, Viswa Anoop N, Midhun PM, Golla Y, Ravi Shankar S, Sathish Kumar P (2013) The auto-dim of vehicle head light using RF-waves. In: Stam JS (ed) Automatic vehicle high-beam headlamp control system, 4th edn, vol 3
5. Karthik MM2016)Automated headlight intensity control and obstacle alerting system,Karnataka, India
6. López A, Hilgenstock J, Busse A, Baldrich R, Lumbreras F, Serrat J (2008) Night time vehicle detection for intelligent headlight control
7. Brick P, Schmid T (2011) Automotive headlamp concepts with low-beam and high-beam out of a single LED, vol 8170. France
8. Narendiranath Babu T, Nakkeeran T, Prasanna V, Rama Prabha D (2021) Auto adjusting brightness of headlights using LDR and ultrasonic sensor to prevent night glare
9. Tamburo R et al (2014) Programmable automotive headlights. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds) Computer vision—ECCV 2014. ECCV 2014. Lecture notes in computer science, vol 8692. Springer, Cham
10. Caroselli, Beachler & Coleman, LLC (2021) What percentage of car accidents are caused by human error

Performance Enhancement of Photo Voltaic System for Rural Electrification in Higher Altitude Region: A Case Study in Uttarakhand, India

Abhiroop Das, R. Aarthi, S. Vijay, M. Kailash, K. S. Gogul Nithish, R. Saravanan, and Souresh Cornet

Abstract One of the greatest challenges in today's world is to satisfy the energy needs of a growing population, in a sustainable way. In Developing countries, actual electrification of rural areas is hindered by high costs and unreliability of service, which often discourage rural communities to avail energy services even when they are available. This study aimed at better understanding the energy struggle of Indian rural communities, and propose efficient technological solutions to address it. Taking the case of a community located in the Himalaya region, a field based, participatory and bottom-up approach is taken to assess the community needs, expectations, and actual resources. A prototype of solar panel using machine learning-assisted automatic sun tracker system is proposed to provide sustainable and affordable electricity to the community.

Keywords Rural electrification · Sustainable development · Rural India · Live-in-Labs®

A. Das
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

R. Aarthi
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

S. Vijay
Department of Chemical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

M. Kailash · K. S. Gogul Nithish · R. Saravanan
Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

S. Cornet (✉)
Amrita School for Sustainable Development, Amrita Vishwa Vidyapeetham, Amritapuri, India
e-mail: soureshcornet@am.amrita.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_38

1 Introduction

One of the greatest challenges in today's world is to satisfy the energy needs of a growing population, underscored by tenets of sustainability. In developing countries like India, access to reliable energy is still a concern. According to the World Bank, even where electricity service is locally available, many Indian village households choose not to adopt a connection [1]. Connection costs are high, and there are no reliable alternatives. One in two rural India grid-users face a power outage of at least 8 h each day [2]. Renewable energies are propitious efficient alternatives that need to be explored. However, the viability of renewable energies to accelerate electrification in remote rural regions faces multiple challenges [3]. In particular, mountainous areas need special attention on technical, financial, and humanitarian grounds to achieve rural electrification. This study conducted as part of the Live-in-Labs® program [4], explores a possible model of self-sustaining rural electrification with the case study of a remote rural community located in the Himalayan state of Uttarakhand, India. Section 2 provides an overview of the study area. Section 3 details the methodology used. Section 4 presents the results of the field work, and Sect. 5 details and discusses the proposed solution.

2 Study Area

Udalka is located in the Uttarkashi district of the northern Indian State of Uttarakhand. Udalka has 121 families in around 70 houses, embedded in a mountainous region with agriculture fields. The houses in this village are located in rugged terrains with limited access to electricity. Dunda, the nearest town, is 6 km away. It has the basic facilities like school, primary healthcare center, markets, and bank [5]. Gadwali and Hindi are the main spoken languages.

3 Methodology

As a key aspect of the Live-in-Labs® program, students and faculty members from multidisciplinary backgrounds are organized into project teams [6]. Four engineering students and two faculty receive an opportunity to experience real-life conditions in the study area, the rural community of Udalka. As live-in guests, the project team resides with the community members for a week to gain a clear understanding of their habitual living conditions. The project team designs a practical solution apposite their findings, which will be appropriate for the needs and expectations of the local residents [7].

The first objective was to understand the energy requirements of the community. Participatory rural appraisal (PRA) was used to derive a comprehensive assessment

of the Udalka community members' energy requirements. PRA is a participatory method used to engage and understand people from their own perspective [8]. PRA is outfitted with a variety of group-based tools, so that all community members can contribute, unrestrained by their age or educational background.

In apposition to the PRA discipline, the project team recorded field notes of all participatory activities and informal discussions with the community. Participant observation was used. Observation was structured around the following areas: activities, environment, interactions, objects, and users, which are part of the AEIOU framework [9]. Quantitative measurement of electricity consumption was also carried out.

The second objective to develop a technical solution with the local residents was achieved with the use of human-centered design (HCD). HCD helps in incorporating the end-users' capabilities and requirements, during the development process of a service or a product [10, 11].

Previous research on developing technological solutions to address socio-economic challenges of rural populations living in mountainous areas [12] have shown that collecting precise information can be an unnerving task. It is advised to triangulate these data with official government statistics to ensure accuracy and validity [13, 14]. Therefore, the research team adopted a mixed method approach, in which primary data collected through HCD, PRA, and participant observation which are largely qualitative, was triangulated with quantitative primary data and secondary quantitative data.

4 Results

4.1 Qualitative Results

The project team stayed with the villagers for one week in June 2019. From the wealth of participatory information gathered using qualitative methods, a summary of the key points is presented in this part.

There is only one solar powered streetlight for the whole village. All the public meetings, festivals, and gatherings are conducted there. From 6:30 pm in summer season and from 5 pm in winter, the community is in complete darkness until the next day. The community members mentioned that the energy supply is unreliable: there are 2 h power cuts per day on average, which can go up to 4 h in the rainy season (July to September).

Household discussions revealed that electricity distribution is limited within the houses. Only the living room and bedrooms, which are considered essential, are electrified. The kitchen and toilets are often left out. Villagers use traditional methods (kerosene lamps, candles, and firewood) to light these rooms, which can have serious health impacts, in particular respiratory issues.

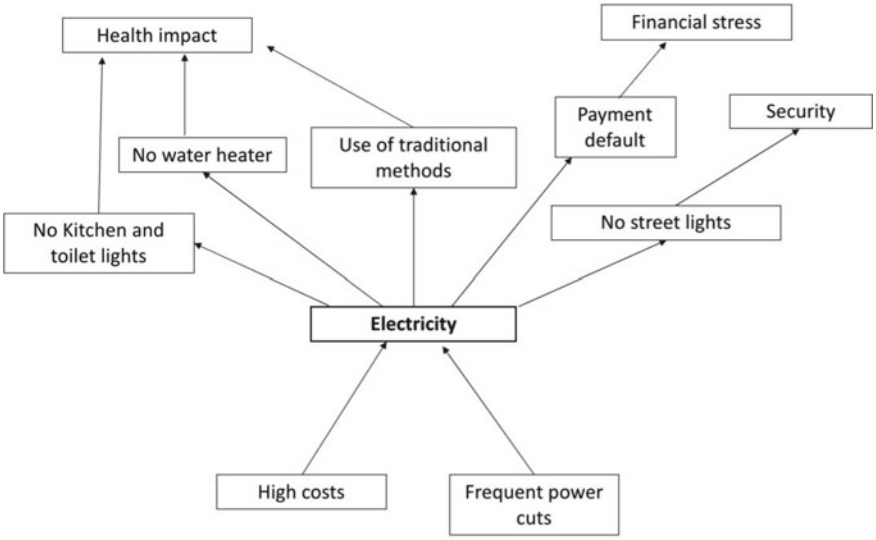


Fig. 1 Problem tree, summarizing the energy challenge faced by Udalka community

Very few families can afford a water heater in winter, when the temperatures go down to 15 °C. This can be attributed to the limited household income compared to high tariffs charged for electricity. The average monthly electricity bill is about Rs. 400 per household, while the average monthly income is about Rs. 2000 per household. The majority of household expenses are allocated for grocery and agriculture (seeds and fertilizers), which does not leave enough to pay for electricity at the end of the month. As a result, households that have not been able to pay their electricity bills have accumulated penalties, further increasing their financial stress.

The challenge of energy is summarized in Fig. 1, using the PRA tool called problem tree [8]. This tool organizes qualitative information related to the challenge of electricity access, with causes (below) and consequences (above).

4.2 Quantitative Results

Qualitative information was cross-checked with the actual energy requirements of the community. The average electricity requirements for a sample of four houses were assessed based on household interviews. As summarized in Table 1, the total requirement for four houses in a day is approximately 36 kW/h.

Meteorological data for the Udalka community were verified with the Indian Meteorological Department (IMD). As shown in Fig. 2—Uttarkashi has an annual average Global Horizontal Irradiance of 5.21 kWh/m^2/day. The maximum solar radiation is available in the month of May and minimum in January. According to secondary data,

Table 1 Daily estimated energy consumption for a cluster of four houses

Appliances	Nb houses	Unit	Watts/unit	Hours/day	Total W/day
LED	4	2	10	6	480
Fans	4	2	55	20	8800
Television	4	1	90	5	1800
Water heater	4	1	1500	2	1200
Refrigerator	4	1	250	24	24,000
Total					36,280

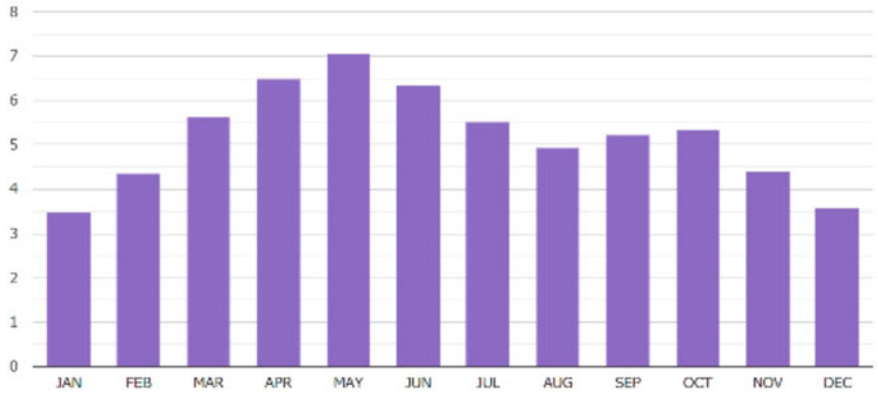


Fig. 2 Monthly average solar radiation data for the Udalka community, 2019. *Source* Indian Meteorological Department

Udalka community has an approximate Global Tilted Irradiation/Irradiance (GTI) of 1978.1 kWh/m^2 per year (for an area of 29,221.35 m^2) [15]. The propitious solar irradiance data indicate that solar power was a promising viable renewable energy in the Udalka community.

5 Discussion

This section presents and discusses the key technical features of the solution to improve energy access in the community. Based on the community's estimated electricity consumption (Table 1), and the solar power data (Fig. 2), the feasibility of solar power as a renewable energy source was systematically investigated in the study area during December–January 2020. The calculation was done with the help of 125 W, 12 V—monocrystalline photovoltaic (PV) panel, along with 12 V, 5 a battery with a charge meter. In the context of the four houses sample (Table 1), a 250-W solar panel produces 1.5 kWh/day. Therefore, 20 solar panels of 250 kW would be required to

Table 2 Electrical parameters specifications of 250 W monocrystalline solar panel based on [16]

Component	Specs.
Power max (*P*m)	250
Short circuit current (*I*sc)	8.95 A
Max power current (*I*mp)	8.35 A
Maximum voltage (*V*mp)	29.95 A
Open circuit voltage (*V*oc)	37.25
System voltage	1000 VDC
Weight	17 kg

supply the Udalka village with 100% of its electricity needs. The footprint needed for set up of the solar panels was approximately 30 m^2. The technical requirements for the solar panels are summarized in Table 2.

The major challenges in high-altitude region is to get consistent solar energy. According to the literature, there are two alternatives to increase the efficiency of solar PV. The first one is concentrated solar PV technique [17]. Solar radiation can be concentrated by positioning the panel appropriately, for generation of the same amount of energy as the conventional system at a lower cost. Simple glass mirrors can be used as reflectors to concentrate the solar energy [18, 19]. However, low investment may be countered by higher training and maintenance costs, in a low-skilled and inaccessible environment.

Another approach is to use a sun tracking system [20]: the PV panel follows the sun during the day and sustains an optimal orientation toward sunlight that maximizes the energy output. Existing sun tracking methods used constant rotation of the PV modules from 9 am to 5 pm [21]. In that model, a sensor controls the panel temperature: whenever it increases above 26 °C; a solenoid valve will be opened [21]. This option seems efficient but requires significant adaptation to the community context.

Based on the models discussed and their respective pros and cons according to the study area, the solution to use sun ray tracker is retained, for maximizing the efficiency of solar panels. However, we propose to enhance it with the use machine learning, as detailed in Fig. 3. With the proposed design, the direction of the sun is continuously monitored by the sun tracker. Irradiance sensors can be used to continuously acquire the irradiance data. Accordingly, the automated PV panel will rotate as per the irradiance of the sun, therefore, enhancing the energy production.

The advantage of using machine learning methods is that it can predict the position of the sun. In this design, it is connected to a motor, that will rotate the panels horizontally. It is recommended to use only the horizontal movement from 30° eastwards to 30° westwards in the terrain region. Horizontal single-axis tracker improves the efficiency by 25–35% compared to a fixed positioned solar panel [22]. Considering the climatic factors of the study area, horizontal single-axis tracker is preferred over other trackers reviewed earlier.

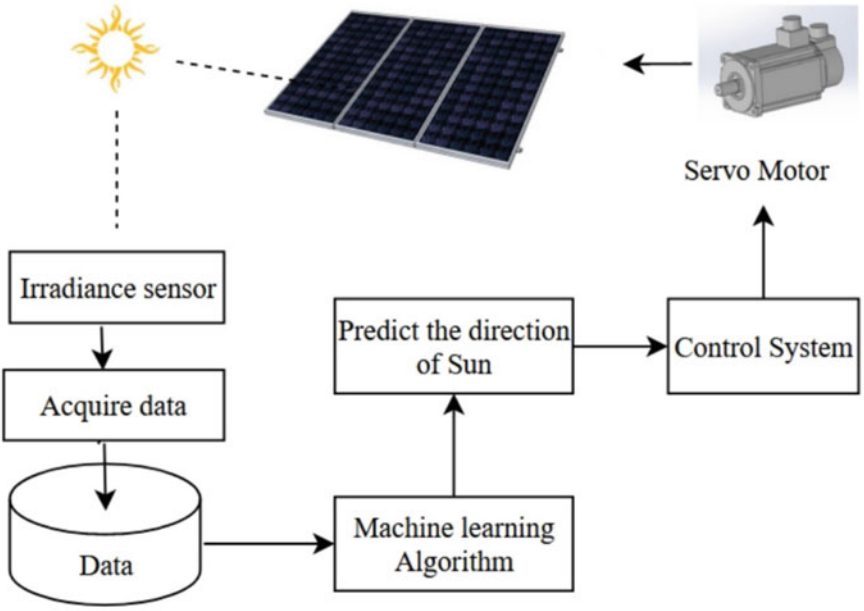


Fig. 3 Proposed design of solar panel system for the Udalka community

6 Conclusion

This study aimed at better understanding the energy struggles of Udalka, a rural community located in India's mountainous regions, in pursuit of a pragmatic technological solution.

The first objective of this study, which was to understand the energy requirements of the community was attained using mixed methods. The energy requirements of the community were estimated based on measurements conducted in a sample of four houses.

The second objective was to develop a technical solution with the local residents. A methodic exploration of Udalka community using mixed methods identified solar power as the most suitable energy option. Using solar energy is in line with the Government of India's objective of augmenting the renewable energy capacity to 175 gigawatts (GW) by 2022, in support of the country's transition to a low-carbon alternative. Of this, 100 GW was earmarked for solar capacity, with 40 GW (40%) expected to be achieved through decentralized rooftop solar projects [23].

The extensive qualitative information collected provided an in-depth understanding of the community members struggles with energy, and therefore, the key requirements for designing a solar-based solution. Pricey capital outlay is one of the main factors that deter rural communities to connect to the grid, even when it is technically feasible [1]. Maintenance and availability of skills locally are other factors that have been discussed.

To cater to the specific energy requirements of the community under study, this paper explored two options for augmenting performance of PV systems. Concentrated solar PV techniques have been proved to reduce the cost of energy generation, while increasing its efficiency [18]. However, it requires maintenance and technical know-how, which are important failure factors in complex settings like mountainous terrains. The other option of using a sun tracker was found more appropriate to the study area [16]. It is a good trade-off for generating sufficient power in a cost effectiveness manner.

Finally, a system design is proposed to adapt sun tracking design to the study area. It uses machine learning for automating the sun tracking process, which reduces the challenges associated with other methods such as cost, maintenance, and skill availability. This is a first level system design proposition. Further, work is ongoing to develop and test a prototype in the study area.

Acknowledgements This study was conducted as part of the Amrita School for Sustainable Development and the UNESCO Chair on Experiential Learning for Sustainable Innovation & Development, at Amrita Vishwa Vidyapeetham. The authors extend their sincere gratitude to the Amrita Live-in-Labs® Academic program for providing all the support. The authors express their immense gratitude to Sri. Mata Amritanandamayi Devi, Chancellor of Amrita Vishwa Vidyapeetham, who has inspired them in performing selfless service to society.

The authors would like to express their sincere appreciation to Mr. Gopalaiengar Sheshadri for his extensive support in enhancing the quality of the paper within the limited time frame.

The authors also thank Subhadraji, the village coordinator, and all the stakeholders of the Udalka community for participating in the study and guiding them throughout the process. Finally, authors thank the government officials of the Uttarkashi district of Uttarakhand, and staff members of the Amrita Self Reliant Villages (Amrita SeRVe) program.

References

1. World Bank. https://openknowledge.worldbank.org/handle/10986/20525. Last accessed 24 Dec 2020
2. Rockefeller Foundation. https://www.rockefellerfoundation.org/report/rural-electrification-india-customer-behaviour-demand/. Last accessed 24 Dec 2020
3. Mohan V, Garcia VC, Frey ML, Ramesh MV, Devidas AR, Shibu NS, Mohan R, Pokkiyarath M (2017) Assessing village power grid problems for development of quality and stable supplemental sustainable energy. In: 2017 IEEE global humanitarian technology conference (GHTC). IEEE, pp 1–10
4. Ramesh MV, Mohan R, Menon S (2016) Live-in-Labs: rapid translational research and implementation-based program for rural development in India. In: 2016 IEEE global humanitarian technology conference (GHTC). IEEE, pp 164–171
5. Government of India. National Census 2011. Last accessed 22 Dec 2020
6. Ramesh MV, Muir A, Nandanan K, Bhavani RR, Mohan R (2022) HCI curricula for sustainable innovation: the humanitarian focus at Amrita Vishwa Vidyapeetham. Interactions 29(1):54–57
7. Varma DS, Nandanan K, Vishakh Raja PC, Soundharajan B, P´erez ML, Sidharth KA, Ramesh MV (2021) Participatory design approach to address water crisis in the village of Karkatta, Jharkhand, India. Technol Forecast Soc Change 172:121002 (2021)
8. Narayanasamy N (2009) Participatory rural appraisal: principles, methods and application. SAGE Publications India

9. Lee MJ, Yuan W, Henry BLD (2012) AR UX design: applying AEIOU to hand-held augmented reality browser. In: 2012 IEEE international symposium on mixed and augmented reality—arts, media, and humanities (lSMAR-AMH). IEEE, pp 99–100
10. Brown T, Wyatt J (2010) Design thinking for social innovation. Dev Outreach 12(1):29–43
11. Ranger BJ, Mantzavinou A (2018) Design thinking in development engineering education: a case study on creating prosthetic and assistive technologies for the developing world. Dev Eng 3:166–174
12. Dey PS, Chettiar MS, Devanand P, Arunajaya CK, Sowmya C, Frey LM, Cornet S (2020) Assessment of sustainable agriculture practices in Uttarakhand, India. In: IEEE 8th R10 humanitarian technology conference (R10-HTC). IEEE, pp 1–6
13. Oberoi A, Mohan H, Basavaraju S, Raparthi S, Bhattacharjee S, Cornet S (2021) Study of an activity tracking device for rural workers through collaborative design. In: IEEE 9th region 10 humanitarian technology conference (R10-HTC). IEEE, pp 1–6
14. Oberoi A, Bandla S, Mohan H, Basavaraju S, Bhattacharjee S, Raparthi S, Frey LM, Cornet S (2022) ICT-based enhancement of employment schemes: a case study in rural Uttarakhand, India. In: ICT with intelligent applications. Springer, Singapore, pp 643–651
15. National Institute of Wind Energy. niwe.res.in/. Last accessed 20 Dec 2020
16. Malik P, Chandel SS (2020) Performance enhancement of multi-crystalline silicon photovoltaic modules using mirror reflectors under Western Himalayan climatic conditions. Renewable Energy 154:966–975
17. Ba¸sak OD, Sazak BS (2014) Effect of components on a solar panel system efficiency. J Res Eng Technol (IJRET) 3:7–11
18. Pavlovi´c ZT, Kosti´c LT (2015) Variation of reflected radiation from all reflectors of a flat plate solar collector during a year. Energy 80:75–84
19. Arshad R, Tariq S, Niaz MU, Jamil M (2014) Improvement in solar panel efficiency using solar concentration by simple mirrors and by cooling. In: 2014 international conference on robotics and emerging allied technologies in engineering (iCREATE). IEEE, pp 292–295
20. Mousazadeh H, Keyhani A, Javadi A, Mobli H, Abrinia K, Sharifi A (2009) A review of principle and sun-tracking methods for maximizing solar systems output. Renew Sustain Energy Rev 13(8):1800–1818
21. Alam SMS, Rahman AM (2016) Performance comparison of mirror reflected solar panel with tracking and cooling. In: 2016 4th international conference on the development in the renewable energy technology (ICDRET). IEEE, pp 1–4
22. Solar trackers: what you need to know, EnergySage. https://news.energysage.com/solar-trackers-everything-need-know. Last accessed 8 Mar 2022
23. Garg V (2019) Vast potential of rooftop solar in India. Institute for Energy Economics and Financial Analysis (IEEFA)

Performance Analysis of Wireless Motes in IoT

Neha Dalwadi and Mamta Padole

Abstract IoT is a network of things or objects that can share information using the Internet. IoT is a set of things connected via the Internet to communicate with each other. As part of wireless sensor network, a set of sensor nodes are known as motes. To accomplish a specific task, different types of motes are existing. In the era of IoT, a wide variety of motes are available for different IoT applications. In the research area of IoT, vast number of motes are available with variety of capabilities. But, due to the scarcity of overall practical knowledge, it is difficult to decide which type of mote should be considered appropriate for a particular application and how to evaluate performance parameters of IoT devices. This paper discusses in details about different types of motes and gives comparative study about the same, for IoT application development. This study will enable one, to identify suitable mote for their application and its performance evaluation. Cooja emulator is used for performance evaluation of different types of motes. It supports a large number of different mote types for IoT applications. The comparative study of specifications is based on motes available in the Cooja emulator.

Keywords Contiki · Cooja emulator · IoT · Mote · Wireless sensor network

1 Introduction

IoT is a collection of things or objects that are connected over the wireless sensor networks and can share information using the Internet. IoT technologies and their applications are growing rapidly by connecting physical objects and by making them smart. With the evolving use of embedded devices, communication technologies, sensor networks, and Internet protocols, it is possible to create smart objects and smart applications. IoT also refers to smart objects that should be small in terms of size,

N. Dalwadi (✉)
Shri K. J. Polytechnic, Bharuch, Gujarat 392001, India
e-mail: neha.dalwadi@gmail.com

M. Padole
The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat 390002, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_39

and low in cost, having effective power utilization, easy to use, and with sufficient wireless range for communication. Hence, it is important to choose an appropriate wireless technology and communication device while building IoT applications. This leads to the proper usage and functioning of sensors and controllers. For example, IoT applications such as armed force inspection require reliable connectivity within the network and may have more power consumption [1]. Whereas in applications, such as environmental monitoring, agriculture monitoring [2, 3] may require less security and may not need full-time connectivity of nodes within the network as compared to defense related applications. In the development of smart applications, wireless motes play an important role in sensing, collecting and communicating, and transmitting data between the motes and data servers. Each mote comprises of tiny sensors, processor or microcontroller, memory, battery for power, RF transceiver, and antenna for communication. Each part of mote has specific role and impact on performance of application. Each of these components has the following purpose and significance.

1.1 Sensor

Sensing unit of a mote consists of different types of sensors depending on an application. A sensor measures external physical properties such as temperature, pressure, humidity, pH value, blood pressure, and heart-beats. These sensed data of physical properties are then passed on to the processing unit, for further analysis. There are two types of sensors available—analog and digital. Sensors have an inadequate power source, and hence, it is important to conserve power in them, for its utility at the appropriate time. Sensors can be enabled or disabled based on requirement, which may impact or conserve power consumption.

1.2 Processor/Controller and Memory

It is core of the mote. It comprises memory, flash storage, analog to digital converter and digital input/output pins. Microcontroller or microprocessor has different level of power consumption, operating voltage, support for external devices and cost. It is important to check these basic characteristics while choosing mote. CPU speed maximizes performance of a processor by minimizing transition time for sleep mode to wake-up mode and vice versa. Depending on an application, and the amount of data generated and required to be analyzed, one should choose CPU speed accordingly. More data analysis and local data processing requires faster CPU.

Mote requires small amount of memory for both program and data storage. Nowadays, microcontroller/processor comes with a flash-based CPU has on-chip storage having less than 128 K memory storage. Data RAM is also available for program execution, and its range is 32–128 KB.

1.3 Battery and Radio Transceiver

Battery is used for power supply to mote. Basically, there are three types of batteries are used in wireless motes: Alkaline—it is economical and high-power source with wide voltage range 1.65–0.8 V. Some sensors do not tolerate high voltage. Lithium—it comes in smaller size than alkaline and works with constant voltage which is better than alkaline battery. One drawback is having low discharge of current unable to drive mote which requires more current. Nickle Metal Hydride—it can be easily recharge. Generate only 1.2 V, and some motes operates on more than 1.2 V [4].

Radio transceiver is a communication unit, responsible for transmitting and receiving data. Low power motes have short-range transceivers. Same power is consumed in data transmission and receiving both directions. If radio is on and whether or not data is transmitted or received, power is consumed.

While developing an IoT application, not only the hardware part is important. One should choose dynamic operating systems to drive hardware platforms efficiently. An operating system that requires less overhead on hardware and runs with low power should be used. There are many real-time operating systems available like Contiki [5], Free RTOS [6], RIOT [7], Tiny OS [8], etc. All these operating systems are compatible with a set of motes to support a low power mode of operations [9]. There are commercially available IoT motes used for general-purpose usage in most applications like MicaZ, Z1 mote, Sky mote, Wismote, ESB mote, etc. All these motes do not support all real-time OS. Hence, the researcher has to choose an appropriate mote with supporting OS in it. Selection of sensor mote for an application is one of the important aspects. As well as one should analyze the performance of developed application based on hardware and software. Due to the complexity of wireless motes and sensor devices, simulators or emulators are used to test or analyzes such applications performance. There is the availability of several simulators and emulators for different types of RTOS so that researcher can test their application before implementing in real time. In this study, Cooja [10] emulator is used to evaluate each mote's performance on Contiki OS. Various types of motes are compared by Cooja based on their features, hardware specification, supported sensors, and power consumption.

2 Contiki OS and COOJA Emulator

Contiki is a lightweight, event-driven real-time operating system. It is open-source software developed by Adam Dunkel's for low power and memory constrained Internet of Things devices [5, 10]. The features supported by Contiki OS are as follows: Pre-emptive multithreading using Protothreads. Protothreads are a concept of stack-less thread; hence, memory is not wasted. Programming language: C, partial C++. There is a wide range of hardware platforms such as AVR, ARM, and MSP430 devices are supported by Contiki. Emulators: Cooja and MSPSim. Both emulators

are based on Java. The code was tested on Cooja, the same code uploaded to motes. The performance of each mote in a network can be analyzed in real time using the collect view tool embedded in Cooja.

Network layer
MAC layer
RDC layer (radio duty cycle)
Radio layer

Protocol stack of Contiki provides communication over uIP which supports IPv4 and IPv6 and Rime network stack. Rime stack supports communication over low power devices with complex abstraction and lightweight communication layer. Network layer protocols CoAP, UDP/TCP, RPL—Routing Protocol, and 6LowPAN. Handles and supports IPv4/IPv6 protocol-based connection and routing mechanism for low power lossy networks. MAC Layer protocols: CSMA/CA designed to handle collision of packets and perform necessary actions like pull back the packets or retransmit. RDC Layer protocol: ContikiMAC allows a node to keep the radio off most of the time for the conservation of remaining power. Radio Layer protocol: IEEE 802.15.4, the lowest layer of the protocol stack that implements the mechanism of IEEE 802.15.4. This layer defines the incoming data structure and store data in packet buffers which are ready to be sent at upper layers in the network.

3 Design Considerations

The Internet of Things demands low power, low cost, and low memory as primary features of devices for the designing of IoT-based applications. There is a trade-off among these three primary requirements of IoT devices. Low-cost devices can be designed by using cheap transistors, semiconductors, crystal, and other parts of mote which can lead to more power consumption. As well as low memory and high-power consumption may degrade the performance of a mote. It is necessary to analyze these factors before the development of new IoT applications to make it efficient and reliable. There are many factors influencing the design of IoT applications. It mainly depends on the hardware and related software in which the application is going to be implemented.

3.1 Power Consumption

One of the major factors while deciding the most suitable mote for specific applications is power consumption. The source of power needs to utilize efficiently in certain applications such as agriculture monitoring, water quality monitoring, and

real-time controlling systems. Power consumption is one of the first design considerations while developing an IoT application that operates on battery power. And hence, power needs to consume efficiently to preserve the residual energy of a node. The most amount of power is consumed during transmitting and receiving messages. Hence, one should choose an energy-efficient communication protocol like ZigBee, Bluetooth, BLE, etc. Power consumption for technology ZigBee/802.15.4 can be determined or calculated by observing the amount of time radio is on or by calculating the current drawn in a circuit. Ideal power consumption is less than power consumption during data transmission or receives. Hence, while choosing a mote, one should know this parameter range and validate that will it be suitable for an application or not. There is a need to know how much power consumption is nominal for our application to run it efficiently. The next section of this paper includes power consumption analysis using the Cooja simulator. Comparison is made for power consumption among different motes available in Cooja for analysis.

3.2 Microcontroller, RF Module, and Life Time of Node

The microcontroller should be able to operate at low power and good processing power. There is a trade-off between power consumption and high processing power. The existing microcontrollers such as the MSP430 series and STM32 series are power efficient that can work with low power and high processing power.

Node lifetime depends on the RF module of a mote. Based on the efficiency of the transmitter and receiver range of the RF module decides the lifetime of a node. The features of RF modules such as transmitter power and range of receivers affect the power consumption of a node. It is very important to identify appropriate RF modules such as CC2420 [11], CC2430 [12], CC2520 [13], RC2400 and related supported protocols to enhance the lifetime of a node. For power-hungry nodes, it is required to make the availability of residual energy at an enormous amount. The module with more power downs and having sleep modes can enhance the lifetime of a node. Table 1 shows some important parameters for the comparison of radio modules from their datasheets [11, 12, 13, 14, 15].

Table 1 Comparison of radio modules

Radio module	RF transceiver	Frequency	Tx	Rx	Tx power
CC2420	IEEE 802.15.4	2.4	17.4	18.8	0
CC2430	IEEE 802.15.4	2.4	27	27	0–5
CC2520	IEEE 802.15.4	2.4	25.8–33.6	18.5	0–5
CC2530	IEEE 802.15.4	2.4	24	29	1
CC1120	Single chip transceiver	164–960	45	17	0–16

3.3 Scalability

Scalability is considered to be one of the factors in terms of communication among a large number of motes as well as in terms of the number of sensors connected and supported by a single mote. Each mote has a limited amount of storage to handle sensor data and processing power for communication; hence, it is required to analyze which type of sensors and how many sensors can be mounted on a single mote such that the mote can work reliably and efficiently.

3.4 Heterogeneity and Interoperability

Heterogeneity and interoperability are some of the important factors, as well as it is one of the challenges in IoT. Research is going on in this direction to resolve interoperability issues among IoT devices, networks, services, and data. Due to incompatibility between IoT platforms, the application becomes platform-specific which can support platform-specific APIs. This will result in a limited accessibility environment of a particular IoT application. It is advisable to check compatibility and interoperability while developing heterogeneous IoT applications. Smart devices or hardware does not support all communication technologies and protocols; hence, there is a need to develop standardization protocols or framework to handle interoperability among devices. There are some solutions available in terms of IoT standards and protocols to handle the heterogeneous nature of the hardware, network, services, and data [10, 16, 17].

3.5 Comparison of IoT Motes in Cooja

Cooja emulator supports different types of wireless motes to emulate IoT applications. It offers a testbed creation for research purposes. Cooja can emulate more than 100 motes and different topology networks. Table 2 shows a comparison of motes based on features and hardware configuration [18, 19, 20, 21, 22] supported by them.

4 Experiment and Results

4.1 Experiment Setup

The purpose of this article is to evaluate the performance of motes which are discussed in Sect. 3. Experiment is performed in Contiki OS with Cooja emulator version 3.0, to examine parameters such as power consumption and memory usage at each mote

Table 2 Comparison of motes based on qualitative parameters

Mote	MicaZ	Sky/TelosB	Z1 (Zolertia)	EXP5438	Wismote
Data rate (kbps)	250	250	250	250	250
ISM band (GHz)	2.4–2.48	2.4	2.4	2.4	2.4
RF transceiver	IEEE 802.15.4 CC2420 radio chip	IEEE 802.15.4 CC2420 radio chip	CC2420 and compliant with IEEE 802.15.4 and ZigBee	CC2500EMK—2.4 GHz radio CC2420/CC2430EMK—2.4 GHz 802.15.4 radio CC2520/CC2530EMK—2.4 GHz 802.15.4 radio	CC2520 and compliant with IEEE 802.15.4 and ZigBee
Processor/microcontroller	MPR2400 (based on Atmega 128L)	RISC processor/MSP430 F1611	RISC processor/MSP430F2617	MSP430F5438	MSP430F5437
Supports	Analog inputs, digital I/O, I2C, SPI, and UART interfaces	SPI, UART, digital I/O ports, watchdog timer, ADC, DAC, supply voltage supervisor, and DMA controller	UART/ADC/SPI	4x UART/LIN/IrDA/SPI 4x I2C/SPI	UART, I2C, Phidgets communication
Hardware security	AES-128	Link layer authentication and encryption	AES-128	–	AES-128
RAM/flash memory	128 kb	10 kb/48 kb	8 kb/92 kb	16 kb/256 kb	16 kb/256 kb
EEPROM (kb)	4	48	96		256
Current draw	8 mA (active mode), < 15 uA (sleep mode)	0.33 mA (active mode)	0.36 mA (active mode), 0.5 uA (sleep mode)	0.312 mA (active mode), 1.8 uA (sleep mode)	0.312 mA (active mode), 1.8 uA (sleep mode)
External power (V)	2.7–3.3	2.7–3.6	2.7–3.3	2.7–3.3	2.7–3.3
Battery	2xAA	2xAA	2xAA or AAA cell	2xAA	2xAA

(continued)

Table 2 (continued)

Mote	MicaZ	Sky/TelosB	Z1 (Zolertia)	EXP5438	Wismote
Integrated sensors	–	Humidity, temperature, and light sensors	Digital temperature sensor and ADXL345 accelerometer	Temperature, humidity, light, or 3D accelerometer	Temperature, humidity, light, or 3D accelerometer
Supported OS	Contiki, Tiny OS, and Mantis	Tiny OS, Mantis, and Contiki	Contiki OS, RIOT OS, OpenWSN, TinyOS, and Mantis	Contiki	Contiki
Data rate (kbps)	250	250	250	250	250
ISM band (GHz)	2.4–2.48	2.4	2.4	2.4	2.4

Table 3 Simulation configuration parameters

Parameters	Associated value	Parameters	Associated value
Topology	Random	Interference range	100 m
Radio medium	UDGM	MAC layer protocol	CSMA ContikiMAC
Mote start-up delay	1000 ms	Channel check rate	8 MHz
Number of motes	3	Network stack	Rime
Radio transmission range	50 m	Mote start-up delay	1000 ms

in a network. Each type of mote is analyzed with the same set of basic parameters. An experiment is done on application included in examples in Contiki OS. Each simulation consists of three motes within transmission range of each other. To ensure reasonable comparison among different types of motes, same application is tested. The set of basic parameters are as follows in Table 3.

This basic setup is done for different types of motes, and powertracer tool of Cooja has been used for power consumption analysis. We have tested power consumption on unicast example. In the directory Contiki/examples of Contiki, there are built in examples available. We need to modify it as per our usage. We have chosen "example_unicast.c" example from "Contiki/examples/rime/" directory to test power consumption at a different stage like power consumption by CPU, at transmission time, and at receiving time of packets. The sample example is slightly modified to measure power consumption. An objective to choose a unicast example is that it includes sending and receive packets to and from nodes in the network. Hence, we can analyze power consumption during the transmission and receiving of packets. Another simple example "hello_world.c" is also tested in which a mote simply prints "hello world" on the terminal. There is no transmission of messages among motes. From this example, power consumption and size of code on each type of mote can be analyzed.

4.2 Results and Discussion

This section shows the results of sky mote, Z1 mote, EXP5438, and Wismote for power consumption and memory usage analysis. Powertrace application gives the result of power consumed during radio on, transmission, and reception of messages in terms of clock ticks. The formula to calculate power (mW) [23] is (Figs. 1 and 2)

$$\text{CPU Power} = (\text{Energest Value} * \text{Current} * \text{Voltage})/\text{Number of ticks per second} * \text{Runtime} \quad (1)$$

where current and voltage are the values as per respective mote datasheet. It refers to current consumption in an active mode of microcontroller and operating voltage. For both the motes, it is the same as shown in Table 2. In our program for time measurement, RTimer is used. The number of ticks per second is 32768. Every 10-s result is taken through powertrace; hence, runtime is 10. The same formula is used to calculate transmission power Tx and Rx, only need to apply current consumption

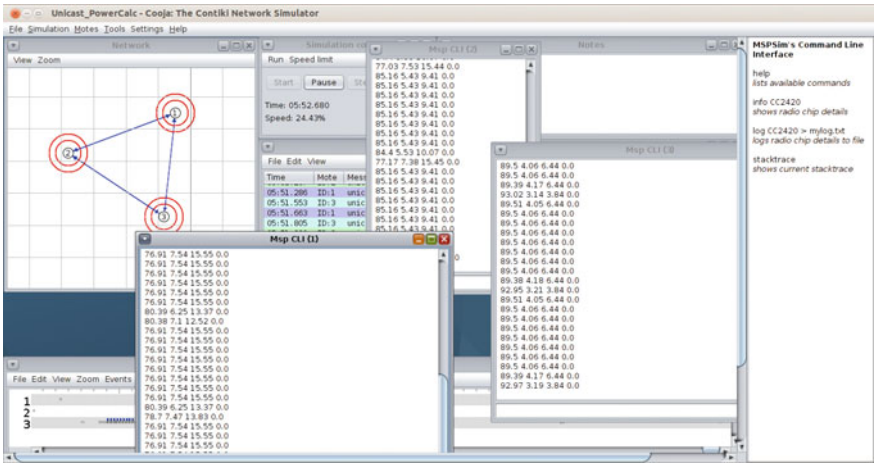


Fig. 1 Screenshot of Cooja with sky motes

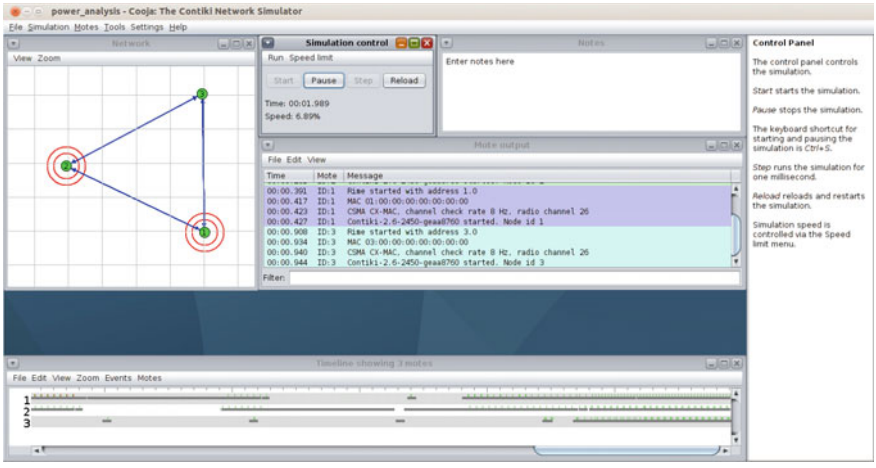


Fig. 2 Screenshot of Cooja with z1 motes

value from the respective platform datasheets [18, 19, 20, 21]. Power consumption is calculated using these values. Results for sky mote are shown in Tables 4, 5 and 6.

Power Consumption Analysis

See Figs. 3, 4, and 5.

Table 4 Sky mote-1 result

CPU (mW)	LPM (mW)	Tx (mW)	Rx (mW)	Total (mW)
0.259256	0.024307	6.15988	5.904031	12.08822
0.260739	0.024257	6.205759	5.937422	12.16744
0.26257	0.024196	6.259444	5.981656	12.2653
0.264093	0.024146	6.304686	6.010745	12.33958
0.247122	0.02471	5.762422	5.628468	11.4156
0.254947	0.02445	6.014279	5.809365	11.84809
0.25241	0.024534	5.930167	5.748091	11.70279
0.258171	0.024343	6.113683	5.878557	12.01658

Table 5 Sky mote-2 result

CPU (mW)	LPM (mW)	Tx (mW)	Rx (mW)	Total (mW)
0.19452	0.02646	4.10170	4.33947	8.46763
0.19455	0.02646	4.10154	4.33224	8.46024
0.19461	0.02646	4.10154	4.34205	8.47005
0.19465	0.02646	4.10218	4.33482	8.46346
0.19468	0.02646	4.10234	4.30143	8.43022
0.19476	0.02645	4.10202	4.30849	8.43696
0.19467	0.02646	4.10218	4.29902	8.42766
0.19471	0.02645	4.10154	4.33620	8.46419

Table 6 Sky mote-3 result

CPU (mW)	LPM (mW)	Tx (mW)	Rx (mW)	Total (mW)
0.14451	0.02813	2.50470	2.95443	5.48725
0.14448	0.02813	2.50422	2.95219	5.48454
0.14454	0.02812	2.50438	2.95597	5.48848
0.14435	0.02813	2.49706	2.95253	5.47772
0.14432	0.02813	2.49706	2.94048	5.46567
0.14438	0.02813	2.49642	2.93979	5.46434
0.14443	0.02813	2.49658	2.94375	5.46846
0.14447	0.02813	2.49690	2.94238	5.46740

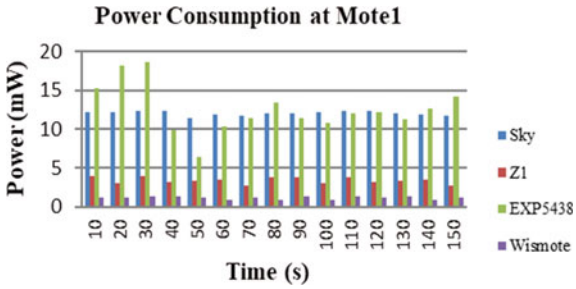


Fig. 3 Total power consumption at mote-1 of different types

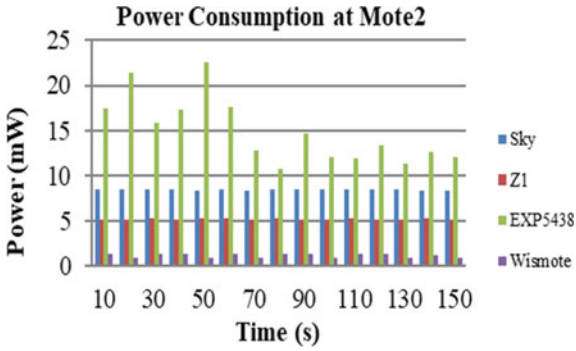


Fig. 4 Total power consumption at mote-2 of different types

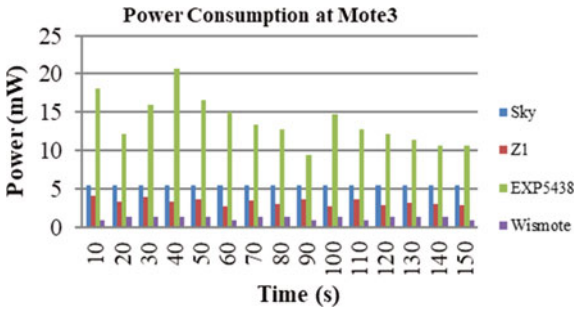


Fig. 5 Total power consumption at mote-3 of different types

Memory Usage Analysis

Memory usage of each type of mote is analyzed based on memory occupied by an application/program in RAM and ROM. Contiki OS facilitates a command to check program and data size occupied in RAM and ROM. Each value in the result is in bytes. The text section shows the size of the program stored in ROM. The data and bss section show the size of memory consumed in RAM. After compilation and generating the output file, the following command is used: > size example-unicast.sky.

Table 7 Memory usage by each mote type

	Sky	Z1	EXP5438	Wismote
ROM (bytes)	23,750	25,380	22,756	19,830
RAM (bytes)	5458	3720	5314	4918

Result:	text	data	bss	dec	hex	filename
	23,750	140	5318	29,208	7218	example-unicast.sky

Table 7 shows the memory usage of different motes.

Performance Analysis

Power consumption is the desired property in our performance analysis of wireless motes. Figure 6 shows the comparison between motes, considering the total power consumption on the same topology. Sky mote consumes two times more power than Z1 and Wismote. The result shows that power consumption by the exp5438 mote is much higher than the other three types of motes. And Wismote consumes less power than the other three types of motes.

Figure 7 shows the standard deviation of the power consumption for the different types of nodes. Figures 3, 4, and 5 show the total power consumption at each node in a network for different type of motes. Here, we can examine that the total power consumed by sky mote and EXP5438 is much higher than other two motes Z1 and Wismote. Performances of Z1 mote and Wismote are efficient than sky mote and EXP5438. All these motes have the same radio transceiver CC2420 but their microcontrollers are different, and hence, there is a difference in the performance of each type of mote. The number of motes in a network can be increased and performance can be evaluated again, but it will be a lengthy process to check the performance of each mote. Also, it can be noted that total CPU power and power consumption during low power mode are almost the same for each type of mote. Hence, this analysis can be useful for the IoT application in which it requires the maximum time of message transmission among nodes in a network. Multiple radio transceivers are supported by various microcontrollers, for example, EXP5438 and Wismote as shown in Table 2.

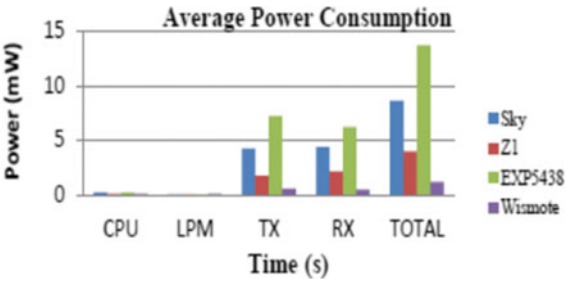


Fig. 6 Average power consumption of different types of motes

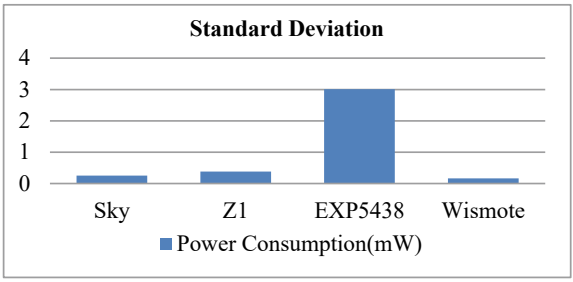


Fig. 7 Standard deviation between different types of mote's power consumption

5 Conclusion

We have presented comparisons of different motes in terms of their energy consumption and memory usage. Among the discussed motes and analysis, it can be concluded that Z1 mote and Wismote are more efficient than the sky and exp5438 motes. Choosing the appropriate radio transceiver and microcontroller combination helps to minimize the power consumption of IoT applications. In the case of memory usage analysis, the Z1 and Wismote have low memory consumption compared to sky and exp5438.

References

1. Ball MG, Qela B, Wesolkowski S (2016) A review of the use of computational intelligence in the design of military surveillance networks. In: Recent advances in computational intelligence in defense and security. Springer International Publishing, pp 663–693
2. Jawad HM, Nordin R, Gharghan SK, Jawad AM, Ismail M (2017) Energy-efficient wireless sensor networks for precision agriculture: a review. MDPI J Sens 1781. https://doi.org/10.3390/s17081781
3. Ingelrest F, Barrenetxea G, Schaefer G, Vetterli M, Couach O, Parlange M (2010) Sensorscope: application-specific sensor network for environmental monitoring. ACM Trans Sens Netw 6(2):17:1–17:32
4. Tuna G, Gungor V (2016) Energy harvesting and battery technologies for powering wireless sensor networks. In: Kolavennu S (ed) Industrial wireless sensor networks. Woodhead Publishing, Sawston, UK, pp 25–38
5. "Contiki Operating System" home page. http://www.contiki-os.org/
6. "Free RTOS Operating System". https://www.freertos.org/
7. "RIOT Operating System". http://www.riot-os.org
8. Levis et al. (2005) TinyOS: an operating system for sensor networks. In: Ambient intelligence. Springer-Verlag, New York, NY, USA, pp 115–148
9. Lajara R, Pelegrí-Sebastiá J, Perez Solano JJ (2010) Power consumption analysis of operating systems for wireless sensor networks.Sensors (Basel)10(6):5809–5826
10. Dunkels, Gronvall B, Voigt T (2004) Contiki—a lightweight and flexible operating system for tiny networked sensors. In: Proceedings of the 29th annual IEEE international conference on local computer networks, pp 455–462
11. CC2420 chipcon products datasheet by Texas Instruments. https://www.ti.com/product/CC2420

12. CC2430 chipcon products datasheet by Texas Instruments. http://www.ti.com/product/CC2430
13. CC2520 chipcon products datasheet by Texas Instruments. http://www.ti.com/product/CC2520
14. CC2530 chipcon products datasheet by Texas Instruments. http://www.ti.com/product/CC2530
15. Lee W, Choi M, Kim N (2012) Different characteristics of radio modules in wireless body sensor network systems. In: 9th international conference on network and parallel computing (NPC), Gwangju, South Korea, pp 308–314, Sept 2012
16. Gambi E, Montanini L, Raffaeli L, Spinsante S, Lambrinos L (2016) Interoperability in IoT infrastructures for enhanced living environments. In: IEEE international Black Sea conference on communications and networking (BlackSeaCom), Varna, pp 1–5
17. Borza PN, Machedon-Pisu M, Hamza-Lup F (2019) Design of wireless sensors for IoT with energy storage and communication channel heterogeneity. MDPI J Sens 19:3364
18. "Micaz datasheet". http://courses.ece.ubc.ca/494/files/MICAz_Datasheet.pdf
19. "Tmote sky datasheet", by Moteiv Corporation. https://insense.cs.st-andrews.ac.uk/files/2013/04/tmote-sky-datasheet.pdf
20. "Z1 datasheet" homepage. http://www.zolertia.com
21. "EXP5438 datasheet", by Texas Instruments". https://www.ti.com/lit/ug/slau263i/slau263i.pdf.
22. "Wismote datasheet". http://www.aragosystems.com
23. Dunkels, Österlind F, Tsiftes N, He Z (2007) Software based on-line energy estimation for sensor nodes. In: Proceedings of the IEEE workshop on embedded networked sensor systems (IEEE Emnets), Cork, Ireland, June 2007

Emotion Prediction of Comments in Twitch.Tv Livestream Environment

Aryan Chouhan, Deep Nanda, Jinit Jain, Kevin Pattni, and Lakshmi Kurup

Abstract Livestreaming platforms are discernibly the most comprehensive sources of data in real time. Such websites enable users to broadcast content like the games which they are playing, while providing them the opportunity to interact with viewers watching the livestream. Twitch.tv is one of the most popular livestreaming platforms across the globe with millions of monthly active streamers and viewers. Owing to the COVID-19 pandemic, there has been a shift in the conventional lifestyle of the people, with them turning towards online alternatives like Twitch.tv for leisure. This change has led to an increase in the engagement of users in these livestreaming platforms by manifolds. Concurrently, a lot of data is generated from this sudden inflow, which can prove very useful in understanding the general consensus of the crowd. This data is very important, and there is a need to construe the true emotion of the people in real time, which is reflected in the comments made by them in the chat section of livestream. The streamers on Twitch.tv can consequently refine their content immediately based on the feedback that they can infer from the responses given by the users. But, due to the sheer volume of data and convoluted nature of the chat due to the use of emojis, emotes, and emoticons, there are bound to be inconsistencies, human errors, and other esoteric references which are exceedingly complex to dissect, making the task of language processing difficult and leading to incoherent results. Taking into account the hindrance posed by these issues, we have taken up the task to achieve fairly accurate emotion prediction by putting forward machine learning and deep learning techniques. This will involve the creation of a labelled dataset that can be used for training and evaluating the algorithms. Given how context-specific most comments are on the platform, this will be an extensive task. The project will also require the creation of an end-to-end system that performs emotion analysis and giving results in real time through feedback-loops.

A. Chouhan · D. Nanda · J. Jain (✉) · K. Pattni · L. Kurup
Dwarkadas J. Sanghvi College of Engineering, Mumbai, India
e-mail: jinitjain2k@gmail.com

L. Kurup
e-mail: lakshmi.kurup@djsce.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_40

Keywords Machine learning · Deep learning · Transfer learning · Emotion analysis

1 Introduction

Natural language processing (NLP) deals with understanding the dynamics of human spoken language and helps computer make inferences. Topic segmentation, machine translation, sentiment analysis, and emotion recognition are some of the most actively researched areas of NLP, and they have shown immense progress over the years. We focus on emotion recognition in a specific context—that of livestreams.

Livestreaming is a very popular vertical that has emerged in parallel to the boom in online gaming in recent times. It involves a 'streamer' broadcasting themselves performing some activity, for example, playing a video game. Today, some of the biggest streamers attract audiences of more than 100,000 live viewers, while some of the biggest streamed events are watched by millions.

Most livestreaming platforms, of which Twitch.tv is the largest, have a live chat enabled that allows any viewer to send any message during the stream. This allows viewers to respond and react to activity occurring on the stream and interact with other users as well. However, with the massive number of people involved, parsing useful information and understanding the audience is an incredibly difficult task. Moreover, the language on Twitch is hyper-contextual and has evolved over time into an amalgamation of memes, emoticons, and esoteric references. Coupled with the volatile nature of the chat, theme extraction and context analysis are extremely difficult tasks—even for humans.

To this end, we propose the use of emotion recognition to extract meaningful insights about the audience. Messages can be classified as 'happy', 'exciting', 'angry', and so on, and using this information an understanding of the audience can be gained. There are different ways of approaching this problem—using a lexica-based approach and using a machine learning-based approach. Specifically, in relation to machine learning, since the task is a classification problem both supervised learning and unsupervised learning methods exist. While unsupervised methods do not require labelled data, both methods require an initial corpus containing structured data.

Since no such corpus exists for live chat data from livestreams, dataset labelling is carried out on a large scale in order to create a corpus on which models can be trained. We identify the emotions of the sentences and label them accordingly. Using this dataset [1], we train supervised learning models such as logistic regression, and decision tree classifiers that predict the emotion of a given sentence.

An adjacent task is that of detecting whether a particular comment is spam or legitimate. Since spam messages can disturb the flow of the conversation and disrupt streams, we wish to identify them and subsequently intimate the user. Since no such corpus exists for this task either, we train machine learning models on Twitter datasets that have a similar data distribution. We then run inference on an unlabelled Twitch dataset and manually attempt to infer how accurate the results we get are.

2 Literature Review

Considerable amount of research has been done on the data generated in Twitch chats and livestreams. In [2], Barbieri et al. analyse the problem of emote prediction, i.e. the task of predicting which emote the user is more likely to use based on a collection of chat room messages, as well as recognizing a certain set of emotes commonly used in troll messages.

Kobs et al. present sentiment analysis based on the Twitch exclusive emotes [3], which can be used by users in the chat function. With the help of a created remote dictionary, they attach the sentiments to the initially unlabelled chat dataset for deriving the dataset used for training their model, which is later fed to a convolutional neural network. Appropriateness of emotes for analysis of sentiment is shown in this paper [3]. In [4], Twitch chat data was analysed to solve an imbalanced binary classification problem of predicting whether the user is a subscriber or not based on comments posted in the Twitch chat section using XGBoost classifier. Reference [5] proposed the creation of a Emoji Sentiment Ranking system which is independent of the European languages for the automated sentiment analysis. Reference [6] aims to understand the emotional changes depicted in the Twitch chat before and after the COVID-19 pandemic. However, one of the shortcomings is that the data used for analysis of the emotions is from only one twitch channel which does not represent the generalized user behaviour. Also, the 'emotes' or emoticons specific to a channel are not considered. In [7], a deep learning-based approach was used to identify offensive language used by the spectators. A transfer learning model, which was trained on Twitter data, was evaluated on data collected from top Twitch gaming livestreams. Although word embedded vectors were not used and only surface text was evaluated to identify offensive comments.

Reference [8] proves that language used on Twitch does not resemble those of other social media because streamers have their personal emotes which have different meanings for their respective channels and [9] takes this research further by explaining the inconsistencies, incorrect grammar, and repetition of the Twitch chats. In order to incite further research into this domain, a case study about the unique features of twitch chat is provided using multiple word vector models. Reference [10] performs sentimental analysis over Twitch data. Machine learning methods are used to classify the chats in the comments section into three classes—positive, negative, and neutral. This does not account for specific user emotions and merely generalizes them into broader classes. Reference [11] is the world's biggest hand labelled dataset of fifty-eight thousand Reddit comments of English language which were categorized into one of 27 emotion classes or neutral. An average $F1$-score of 0.46 is achieved on BERT-based model across their proposed method, indicating that there is still significant opportunity for improvement.

3 Methodology

3.1 Dataset Description and Labelling

Out of approximately 3 billion unlabelled chat messages made available by [3], which comprised of the comments posted by users in the most active English Twitch.tv channels, around 5000 were randomly sampled to prepare a labelled corpus for training the model. This dataset includes timestamp of the comment, the streamer whose stream on which the comment was posted, the current game being played, whether the user is a moderator or not, whether the user is subscribed to the streamer or not, and the comment posted. Only the message was retained from this data.

The comments were categorized into seven different emotions—happiness, sadness, anger, surprise, disgust, excited, and neutral. This is based on [12] which identified six basic human emotions—enjoyment, sadness, anger, surprise, disgust, and fear. Based on direct observation of activity on Twitch.tv, fear was excluded, excited, and neutral were introduced, and enjoyment was replaced with happy. Excited refers to extreme happiness and can be differentiated from enjoyment on the basis of the intensity of the message. Neutral refers to messages with no perceivable emotion. This was done because interactions on Twitch were found to be significantly likely to invoke these emotions, hence making them relevant for the streamer to perceive.

A big challenge in labelling comes from the use of emotes on Twitch. Emotes differ from traditional emojis and emoticons in that they are specific to Twitch. Global Emotes are usable across all streams on Twitch, and Custom Emotes are streamer-specific. The relatively obscure meanings of each emote and their variation across contexts make them especially difficult to analyse for emotion.

From a data labelling point of view, each comment was labelled by taking into account the overall meaning of the sentence combined with the emotion evoked by the emotes (if present in the message). The data labelling process is subject to various biases, including inherent biases in Twitch user messages, or inherent bias in assessment of abusive identity labels. Moreover, the annotators were all fluent English speakers from India with reasonable familiarity with Twitch language. All these issues likely impact labelling accuracy and propagate into the trained models too. Any use of this dataset should account for these limitations (Fig. 1).

3.2 Pre-processing

The data is cleaned first using pre-processing techniques, after which, the entire corpus of data is divided into three parts, i.e. training, testing, and validation dataset. Next step is to extract features using various feature extraction techniques, and then lastly, the model is trained and evaluated.

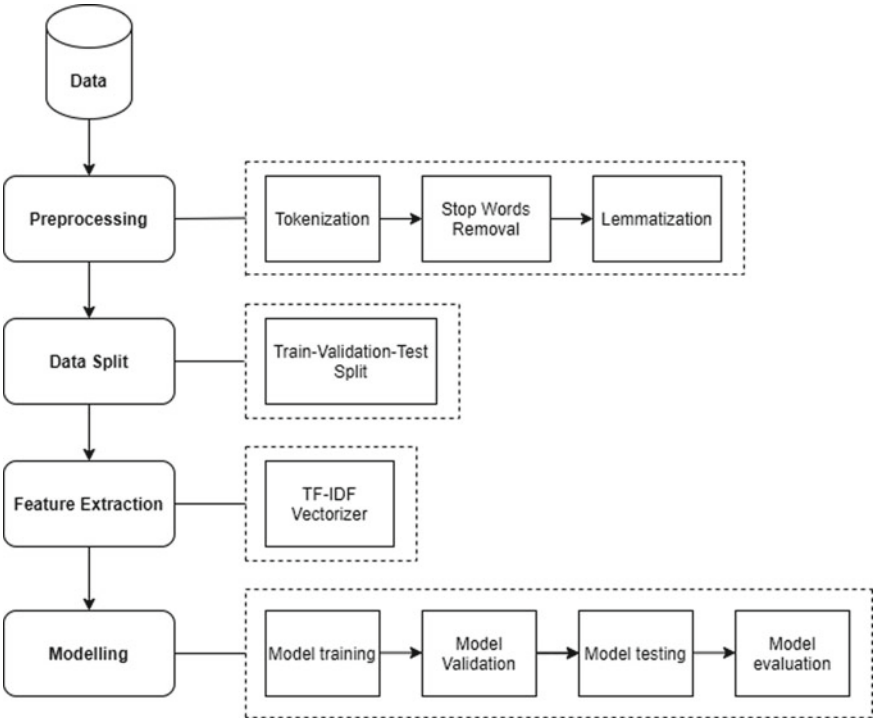


Fig. 1 Process workflow

Tokenization. It follows forming tokens, in simpler terms, it is the process of breaking a piece of text (i.e., a sentence) into smaller units called tokens (i.e., words). The method, same as that used in [3], was implemented for basic pre-processing and tokenization. This was used for text data which may include plain text, hyperlinks, hashtags, numbers, emoticons, and twitch emotes. The data, in the form of a comment, is tokenized and converted to lowercase while preserving the emoticons. The identification of the emotes was done by mapping them to the lexicon created by [3] that consisted of top 100 frequently used emotes. Lastly, 'URL' was used to replace the occurrences of hyperlinks and for characters that occur more than twice in succession were capped to two occurrences. *Sample Heading* (*forth level*). The contribution should contain no more than four levels of headings. The following these methods are primarily used to generate the root form of the words. Stemming involves removal of suffix from the word while lemmatization involves the morphological analysis, i.e. considering the context and then converting the word to its root form. Lemmatization of the data was done using WordNetLemmatizer [13], based upon Princeton's WordNet [14, 15] gives a summary of all heading levels.

Stop Words. Stop words are mostly commonly occurring words that add no significant meaning to any data. Thus, these redundant words are removed to reduce the size of the dataset and obtain better performance. To perform this task in accordance

with the proposed system, NLTK's [13] already available stop word list is used with the modification implemented in [3]. The modification is—those stop words that contain apostrophes were added after removing the apostrophe. This modification is adopted to handle the common case of misspelling the word while commenting in a livestreaming environment.

Lemmatization and/or Stemming. These methods are primarily used to generate the root form of the words. Stemming involves removal of suffix from the word, while lemmatization involves the morphological analysis, i.e. considering the context and then converting the word to its root form. Lemmatization of the data was done using WordNetLemmatizer [13], based upon Princeton's WordNet [14, 15].

3.3 Feature Extraction

The tokenized data must be vectorized as the data expected for the model building is in the form of tensors of numeric values. The dataset has been broken down into 60%–20%–20% for the purpose of training, hyperparameter tuning, and evaluation.

Count Vectorization. Count vectorization involves counting the frequency of each token present in a document. These counts are then used to encode the tokens.

Term Frequency—Inverse Document Frequency (TF-IDF) Vectorization. This vectorization technique uses a TDF-IDF transformer in order to normalize the count obtained from count vectorization. The purpose of normalizing the count is to provide weights to each token which define the relevance of the token in the document. This method to re-weight the count helps to prevent bias towards high frequency tokens. TF-IDF vectorization was used for feature extraction as it determines the relevance of a token, thereby efficiently improves the performance by preferring only relevant data.

3.4 Modelling

Different machine learning and deep learning classification techniques were used to predict the emotion of the comment.

The machine learning models used are—support vector classifier, decision tree classifier, logistic regression, random forest classifier, and multinomial Naïve Bayes. The deep learning models used are—BERT [16], RoBERTa [17], and XLNet [18]. For these, feature extraction is skipped. Transfer learning was performed by fine-tuning the pre-trained BERT-based deep learning model which was trained on GoEmotions dataset [11].

Table 1 Comparative analysis

Model	Accuracy	Macro average precision	Macro average recall	Macro average $F1$-score
BERT	55.0%	41.9%	44.3%	42.5%
Logistic regression	52.3%	60.5%	35.0%	38.7%
Random forest classifier	51.5%	54.8%	36.3%	39.6%
RoBERTa	50.3%	41.6%	45.2%	41.4%
Support vector classifier	49.7%	48.5%	35.6%	38.2%
Decision tree classifier	48.2%	45.7%	36.1%	38.5%
Multinomial Naïve Bayes	48.1%	64.5%	26.8%	27.2%

4 Results

4.1 Model Evaluation

The performance of each model was measured against the test set. The metrics—macro average $F1$-score accuracy, macro average recall, macro average precision, and accuracy of various models were chosen for comparative analysis. The results are displayed in Table 1.

The macro average was used instead of the weighted average since the dataset is imbalanced. If the weighted average is used, a class with fewer labels has less impact on the result. The macro average negates this imbalance by giving each class equal importance. From these results, it can be observed that BERT model outperforms the rest of the models.

4.2 Transfer Learning

Pre-trained model was fine-tuned on the Twitch.tv labelled dataset [1] using a lower learning rate with variation in batch size to fit the best according to the dataset available. The accuracy of 58.05%, precision of 55.99%, recall of 58.05% and $F1$-score of 55.01% was obtained by this approach.

5 Conclusion and Future Work

Through this work, we have applied various machine learning and deep learning techniques to perform emotion analysis on comments posted on Twitch.tv. In order to perform emotion analysis, we have manually labelled the dataset and classified the comments with one of the seven emotions that best fit the comment. By observing the dataset thoroughly, we understand that the importance of the knowledge of emotes is paramount in order to make sense out of the esoteric and dynamic language of not just Twitch.tv, but livestreaming platforms in general. Using the methods applied by us, the streamers can get a better understanding of the responses of the viewers and improve their stream accordingly.

There is a lot of room for expansion in our work, especially when considering the boom in the e-sports competitions and their consequent streaming. We can modify the models further by integrating more streamer-specific emotes and general twitch emotes in it. The accuracy of our methods can be enhanced more with further research.

References

1. Dataset. https://github.com/deepnanda30/TwitchEmotionAnalysis/tree/master/Dataset
2. Barbieri F, Espinosa Anke L, Ballesteros M, Soler J, Saggion H (2017) Towards the understanding of gaming audiences by modelling Twitch emotes. In: Proceedings of the 3rd workshop on noisy user-generated text, pp 11–20, Jan 2017. https://doi.org/10.18653/v1/W17-4402.Author F, Author S (2016) Title of a proceedings paper. In: Editor F, Editor S (eds) Conference 2016, LNCS, vol 9999. Springer, Heidelberg, pp 1–13
3. Kobs K, Zehe A, Bernstetter A, Chibane J, Pfister J, Tritscher J, Hotho A (2020) Emote-controlled: obtaining implicit viewer feedback through emote-based sentiment analysis on comments of popular Twitch.tv channels. ACM Trans Soc Comput 3:1–34. https://doi.org/10.1145/3365523. Author F (2010) Contribution title. In: 9th international proceedings on proceedings. Publisher, Location, pp 1–2
4. Gaertner M, Theissler A, Fernandes M (2020) Detecting potential subscribers on Twitch: a text mining approach with XGBoost—discovery challenge ChAT: CoolStoryBob
5. Kralj Novak P, Smailović J, Sluban B, Mozetič I (Dec 2015) Sentiment of emojis. PLoS One 10(12):e0144296. https://doi.org/10.1371/journal.pone.0144296
6. Chae SW, Lee SH (2022) Sharing emotion while spectating video game play: exploring Twitch users' emotional change after the outbreak of the COVID-19 pandemic
7. Gao Z, Yada S, Wakamiya S, Aramaki E (2020) Offensive language detection on video live streaming chat. 1936–1940. https://doi.org/10.18653/v1/2020.coling-main.175
8. Olejniczak J (2015) A linguistic study of language variety used on Twitch TV: descriptive and corpus-based approaches, 21–23 May 2012
9. Ringer C, Nicolaou M, Walker J (2020) TwitchChat: a dataset for exploring livestream chat
10. Chouhan A, Halgekar A, Rao A, Khankhoje D, Narvekar M (2021) Sentiment analysis of Twitch.tv livestream messages using machine learning methods. In: 2021 fourth international conference on electrical, computer and communication technologies (ICECCT), pp 1–5. https://doi.org/10.1109/ICECCT52121.2021.9616932
11. Demszky D, Movshovitz-Attias D, Ko J, Cowen A, Nemade G, Ravi S (2020) GoEmotions: a dataset of fine-grained emotions. In: Proceedings of the 58th annual meeting of the association for computational linguistics, pp 4040–4054

12. Ekman P (1999) Basic emotions. In: Dalgleish T, Power T (ed) The handbook of cognition and emotion, pp 45–60
13. Bird S, Loper E, Klein E (2009) Natural language processing with Python. O'Reilly Media Inc.
14. Miller GA (1995) WordNet: a lexical database for English. Commun ACM 38:39–41. https://doi.org/10.1145/219717.219748
15. Fellbaum C (1998) WordNet: An Electronic Lexical Database. MIT Press, Cambridge, MA
16. Devlin J, Chang M-W, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. Cite arxiv: 1810.04805. Comment: 13 p
17. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) RoBERTa: a robustly optimized BERT pretraining approach. Cite arxiv: 1907.11692
18. Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov R, Le QV (2019) XLNet: Generalized autoregressive pretraining for language understanding. Cite arxiv: 1906.08237. Comment: Pretrained models and code are available at https://github.com/zihangdai/xlnet

A Review of Deep Learning Healthcare Problems and Protection Supports

D. Karthika and M. Deepika

Abstract Due to the increase in many tools, the relevance of data analytics in health data systems has been growing tremendously as a result of this exponential increase. As a result, it has also led to renewed confidence in the development of data-oriented frameworks for sensor learning in health information systems. With funding provided by artificial neural networks, researchers are looking into the possibility of creating a flexible machine that redefines the recent vision of artificial intelligence. As computing power and rapid processing of reality have increased, the rapid return on technology has often been compromised in terms of conceptual control and production of highly functional and semi-relevant input details that are mechanically designed tails. In this paper, we provide an in-depth analysis of the relative relevance and potential downsides of the approach and its attitude as well as an advanced examination of fundamental learning in health information systems. Additionally, the paper discusses the challenges, security, and protection in healthcare systems with the use of machine learning approaches.

Keywords Deep learning · Machine learning · Health informatics · Bioinformatics · Clinical imaging · Natural language processing

1 Introduction

In comparison with other sectors, the healthcare segment is unique. It is a high-needs area, and people expected the highest quality of care and services, even though it has taken a toll. Although consuming a massive portion of the budget [1], the restorative master is usually the one who does the translations of restorative material. Due to its subjectivity, the intricacy of the picture, and the wide variety of translations and

D. Karthika (✉)
Patrician College of Arts & Science, Adyar, Chennai, Tamil Nadu, India
e-mail: drkarthika@patriciancollege.ac.in

M. Deepika
B. S. Abdur Rahman Crescent Institute of Science & Technology, Chennai, India
e-mail: deepika.ca@crescent.education

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_41

fatigue, the human master's ability to elucidate the picture is severely limited. Deep learning has offered invigorating arrangements with great precision of therapeutic imaging and is viewed as a crucial strategy for future applications in the wellness sector, after its success in other real-world applications. The days of healthcare data being tiny are long gone. In conventional neural arrangement cases, we can quickly demonstrate the computational premise of more profound instruction. As opposed to the more conventional neural arrangement strategy, deep learning uses more layers and neurons [2]. A wide variety of neurons ensures a comprehensive range of non-processed fats; however, a strong nonlinear sorting process reduces the size of the inclusion area. A projection with a lower measurement is designed to create a significant amount of energy. In an optimally weighted organization, unwholesome information or images are precisely envisioned [3].

In addition to helping with the identification and extraction of highlights, deep learning will assist in the advancement of new ones [4]. As well as evaluating the ailment, it will also provide relevant figures to help specialists be more effective. Machine learning (ML) and artificial intelligence (AI) have come a long way in the past few decades. Computer-aided picture treatment, computer-aided conclusion, picture explanation, picture fusion, picture join, picture division, image-guided treatment, picture recovery, and inquiry have all benefited from ML and AI techniques. AML strategies effectively extract information from pictures and converse with them.

DL computations have been consumed, although a few subsequent surveys have led to an increasing dread of machine security and constraint [5]. So also, a few sorts of information and demonstration of harmful attacks against DL methods have been expected, as well as various countermeasures. In any case, the security mechanisms' productivity is obscure, and various ponders have demonstrated that most assurance strategies come up short to anticipate an assault. The truth is that DL frameworks don't have a sound but practical application for unsafe restorative occupations like life-predictive operations. The sector's rigid application of ML/DL strategies is additionally basic. This paper contains an in-depth study on the relative relevance and potential downsides of the approach and its attitude and an advanced examination of fundamental learning of health information systems. It also discusses the challenges, security, and protection in healthcare systems using machine learning approaches.

2 Literature Survey

The analysis, monitoring, and management of human activities in health care can be accomplished using a variety of approaches. There are different solutions that they offer. In a systematic literature review, Qureshi et al. [1] investigated whether ambient supported living could motivate patients with heart disease and enable them to self-manage their condition. The objective was to minimize the mortality and morbidity associated with their disease. Studies include self-monitoring, wearable systems, clinical management systems, ambient assisted living in elderly populations,

and deep learning-based diagnosis of heart disease. In their first report, Motwani et al. [2] describe a comprehensive survey on ubiquitous networks and smart healthcare systems that can be used to monitor chronic diseases and lifestyle diseases. This approach combines cloud-based analytics and deep learning to provide a smart patient monitoring and recommendation system. A case study was presented on the imbalanced dataset gathered on a patient with chronic hypertension, and the status of the patient was discussed. According to the experimental data, the approach was effective. Deep learning applications in health care have been described by Yao et al. [3]. Seven areas of deep learning were focused on. These include electrocardiography, electronic health records, community health care, electroencephalograms, genomics, and drug analysis, as well as data from wearable devices. There is an analysis of the advantages and disadvantages of the existing studies as well as a discussion of challenges and future trends.

3 Methods: ML for Healthcare Problems and Its Challenges

There are numerous obstacles to ML/DL systems being implemented in real-world healthcare settings [4].

1. **Safety Challenges**: Experiential success in a research facility setting doesn't equate to security. The secure ML/DL process ensures that patients are properly cared for inside the ML/DL process. Patients must be assured consistently throughout the process. There is a routine job for the specialist, and the patients are at distinct stages of well-being [6]. Disarrays of wellness that occur every million years must be distinguished from unusual, helpless, and perplexing events. ML/DL transmission to ensure the security of contemporary AI systems in unexpected strata, exceptions, boundaries, and helpless conditions.
2. **Privacy Challenges**: Patients believe they are complying with fundamental assurance conventions through their wellbeing controllers to preserve their unique rights to protection, such as age, sex, date of birth, and health data [7]. Potential security dangers may be of two sorts, uncovering touchy subtle elements and unfeeling utilization of information. Protection depends on the exact setting and nature of the truths, the climate, and the profiles of shoppers. Moved forward protection manhandle utilizing worthy strategies is moreover hazardous since such infringement will harm the customer straightforwardly [8]. To prevent encroachment of protection, such as re-identification of individuals, security ought to be explored. Too, it is vital to consider protection issues at all levels of information collection and to guarantee that information is reallocated between different units inside a clinic in a secured circumstance.
3. **Ethical Challenges**: Sometimes recently gathering realities on the organizing of ML structures, particular measures must be taken to distinguish the aim of patients and their sociological stages. Too, it is the most thought to concur that

the gathering of proof will influence the well-being and self-esteem of a patient [9]. If ethical issues are not detailed, it is troublesome to discover findings within the capacities of ML strategies in operation. Moreover, it is imperative to supply a great visual see of the AI component in eccentric and complicated circumstances to ensure the rise to and ethical work of automated frameworks.

4. **Causality is Challenging**: Causality is the foremost critical component in well-being treatment since most basic wellbeing conditions include fundamental consideration. Conventional learning calculations cannot control such issues and must analyze the evidence from causal contacts to reply to them. Mindfulness, which is, for the most part, centered on assessment truths, is frequently very troublesome in health care since it needs the creation of causal models to address fundamental questions. DL structures require principal presumptions, and these models work fundamentally without an essential association by leveraging tests and relationships. In certain conditions, this cannot be treated as a sickness, since the forecast includes no fundamental relationship. Decision-making will at that point be compulsory through interaction through causal investigation. It is necessary to gauge the causal effect of any information on the objective generation to preserve reasonable expectations [10].
5. **Policy and Regulatory Challenges**: The full potential of the ML/DL systems for addressing administrative and approach issues under ordinary wellbeing conditions will be realized eventually. This administrative direction is required both for ML/DL restorative approaches and their incorporation into standard restorative settings [11]. It should be taken into consideration that the integration of AI-enabled ML/DL frameworks should also follow government and regulatory guidelines. A huge issue with the utilization of therapeutic ML/DL calculations is how those frameworks can be run and controlled since these frameworks execute enhanced user information learning. Accessible controls are lacking for certifying ever-changing items, such as ML/DL frameworks. Further, it is vital that establish, viable, and steady convention is preserved for the therapeutic examination of the ML/DL component in therapeutic conditions that do not compromise the understanding [9]. Data scientists and AI engineers in hospitals can ensure AI is protected, relevant, and functional by regularly reviewing its methods.
6. **Availability of Good Quality Data**: The arrangement of high-excellence prove may be a noteworthy wellbeing concern such as since, relative to heterogeneous collections of wide multi-systems, the sum of data available to the inspected populace is inconceivably restricted and competent. Distinctive little and enormous wellbeing organizations create therapeutic records every day [7]. The development of high-quality truths, closely resembling medical conditions, is unimaginably troublesome and ought to be facilitated and secured. The openness of high-quality actualities will back the potential work of scourge discovery and planning decision-making viably. The information compiled endure from different issues counting subjectivity, excess, and segregation. As ML/DL models will derive from the mystery factors of the information they are qualified for. Subsequently, the impacts of the calculation would be imitated by the information created by the undesirable past involvement in healing centers. Many people without health

insurance are denied access to emergency rooms, and if AI is aware of this, it can do the same. The machine might explain racial bias by producing erroneous findings for diverse subpopulations, and it can also cause modeling errors.

7. **Lack of Data Standardization and Exchange**: The ML/DL system will aid create a profound understanding of straightforward sterile obligations through the utilization of other shopper subtle elements. Radiology, for illustration, does not cover imaging procedures. For radiologists, another EMR reality is essential if an imaging investigation is to realize an exact conclusion [12]. This incorporates meeting and sharing information on overall therapeutic strategies. Noticing that the taking after criteria of IT wellbeing administrations, regularly influencing the greatness and viability of healthcare truths, has been commonly disregarded, during thinking about common concepts of information sharing on open well-being. There are numerous wellbeing strategy conventions such as imaging trials that keep up the therapeutic effect of the proof. Display IT programs for wellbeing care disregard fundamental values and doctors are not embracing well-known directions. The integration of data and the exchange of efforts across areas and cultures thus becomes a challenge. The integration of data into numerous users' health records is critical for patient care. The lack of interest in the wider health industry for data-trading concepts slows the efficacy of ML/DL, as multi-modal data are required to enable a deep understanding of algorithms and to improve clinicians' performance in making therapeutic choices based on data-driven insights.
8. **Distribution Shifts**: The complexity of information sharing is likely the greatest deterrent and the most challenging issue to deal with. For a variety of reasons, such as therapeutic honing, educating, and investigating, the dissemination of proving may be overlooked. Due to this problem, ML/DL programs built using open datasets in actual healthcare conditions do not achieve the results expected. In a medical setting, clinical creativity, in which different conventions and obligations can impact pictures of distinctive dispersions, is commonplace [10]. Model learning frameworks are qualified primarily by the experimental hazard minimization hypothesis, which sets limits on their learning and determines if they are effective. One of the key and most effective claims is that the planning and request information are derived from different sources. The statement isn't sufficient, however, frameworks based on this assertion shouldn't generalize extra spaces. However, this declaration is insufficient, and the systems taught under it should not generalize to other domains. Comparatively, medical applications demand reliable and secure ML/DL devices.

4 Security, Privacy, and Robust Machine Learning for Health Care

This section discusses the protection and robustness of ML/DL under health circumstances, as well as other related problems [13]. In addition, three critical phases of

the ML model's growth are identified, as well as a slew of potential weakness sources that represent significant hurdles at every stage of the ML pipeline.

The protection of ML: A broad notion here gives a general sense of ML safety from the standpoint of health care, and the application of the ML system in [12] raises several associated safety issues.

Data Collection Vulnerabilities: Planning of ML/DL models for restorative appraisal back needs a bunch of expansive amounts of information which is ordinarily time-consuming and needs major human exertion. Indeed, although in hone, clinical information is primarily collected to create beyond any doubt the productivity of the malady. Still, there are so numerous causes for vulnerabilities that can influence the right utility of the basic ML/DL structures. The composed measurements by and large incorporate a parcel of artifacts that rises because of two instrumental and natural flimsiness. This methodology is amazingly touchy to developments, even a little development of the subject's head or breath may lead to undesirable artifacts within the auxiliary picture, in this manner raising the hazard of misdiagnosis. A healthcare environment is exceptionally intriguing and contains specialized and non-technical human assets and a need for gifted laborers that can grow and maintain ML/DL frameworks. Employees with significant theoretical and computational understanding, such as engineers and data scientists, are required for the well-organized function of data-driven health care. The medical usability of ML/DL foundation systems, on the other hand, is critical.

Vulnerabilities Due to Data Annotation: Most of the ML/DL capacities in healthcare frameworks are directed ML activities that want a large amount of labeled data. The method of assigning names to every shown information is known as information explanation. Ideally, this activity will be carried out by competent individuals who will organize domain-enriched datasets that are critical for the creation of ML/DL models in healthcare settings. This collection has revealed that arranging ML/DL models without a solid understanding of the area may be ineffective. Although, specialists like master radiologists are constrained experts and intense to delegate in auxiliary occupations like information explanation. In this way, disciple staff with slight space information are by and large utilized amid information labeling, which for the most part leads to a few inconveniences, for occurrence, coarse-grained names, lesson imbalance, name outpouring, and disarrangements. Several exact information annotation-based vulnerabilities are being examined.

1. Model training vulnerabilities: Unsuitable, insufficient training, invasion of privacy, process corruption, and piracy are all examples of model planning weaknesses. When the ML/DL method's credentials are incorrect, conditional preparation is wrong or poor. Furthermore, ML/DL is particularly specified as being vulnerable to a variety of security and privacy threats, including unfavorable assaults, device attacks, and data corruption attacks. In addition to its presence, the strain of ML/DL processes retains its technological usage for security-critical operations and essential applications. As a result, ensuring the security and efficiency of the ML/DL approaches for such important applications is crucial.

2. Implementation vulnerabilities: The application of machine learning and deep learning in medicine necessitates a human-centered approach. It is consequently critical to preserve the framework's robustness in terms of fairness and openness during the implementation stage. The following flaw can be addressed as part of the ML/DL implementation process.
   (a) Distribution changes: In basic healthcare situations, conveyance shifts are gigantically anticipated, for case, different imaging centers, and afterward, DL frameworks qualified for pictures of one range are utilized on numerous field pictures. The basic DL gadget is altogether corrupted in these conditions. In comparison, prescient ML programs are built from past quiet information and often involvement the utility of the ML arrangement in unused patients. Such a difference may too be demoralized for antagonistic advancement.
   (b) Incomplete data: Missed remarks or variations in genuine surroundings can be found in facts compiled for the patient's benefit. On the other side, using missing ML/DL method values adds to two well-known issues: false positives and false negatives. Both concerns would have considerable repercussions in real-world healthcare settings. To provide accurate result estimates, the clinical information should be thorough and compacted.
3. Measuring process vulnerability: Testing stage helplessness influence the investigation of the introductory ML/DL framework result. False-positive and false-negative results, error. Dishonorable frameworks preparing or somewhat completed information bolstered within the finding recorded within the going before the portion is the cause for untrue positive and untrue negative outcomes. At last, not only turns a wrench into the precise nature of ML-enabled health care but emphasizes the cautious position of expository forms as well.

5 Conclusion

Pattern recognition and digital culture have improved as a result of deep learning. This paper discusses how data-driven applications, utilizing digital technologies, have helped to drive the development of health informatics. Advanced technologies such as diagnostic tests, medical equipment, and bioinformatics have been developed as developed. The region offered an interesting machine learning pattern, but deep learning should not be considered for each test offered by healthcare IT. As a result of fast learning algorithms capable of producing consistent results with little energy use, little parameterization, little tuning, and advanced interpretability, it has been argued that extensive training expertise and computing power are required for optimum performance. As a result of accepting the current advantages of simultaneous treatment enabled by co-processors, the hoped-for revolution in NNs and networking has occurred. However, developing new machine learning algorithms

that have increased computational performance and interpretability requires time and effort.

References

1. Qureshi MA, Qureshi KN, Jeon G, Piccialli F (2021) Deep learning-based ambient assisted living for self-management of cardiovascular conditions. Neural Comput Appl
2. Motwani A, Shukla PK, Pawar M (2021) Novel framework based on deep learning and cloud analytics for smart patient monitoring and recommendation (SPMR). J Ambient Intell Humanized Comput
3. Yao Z-J, Bi J, Chen Y-X (2018) Applying deep learning to individual and community health monitoring data: a survey. Int J Autom Comput 15(6):643–655; Chen Y-W, Jain LC (2020) Deep learning in healthcare. Springer, Berlin
4. Qayyum A, Qadir J, Bilal M, Al-Fuqaha A (2020) Secure and robust machine learning for healthcare: a survey. arXiv preprint arXiv:2001.08103
5. Muhammad K, Khan S, Del Ser J, de Albuquerque VHC (2020) Deep learning for multigrade brain tumor classification in smart healthcare systems: a prospective survey. IEEE Trans Neural Netw Learn Syst
6. Zhou X, Liang W, Kevin I, Wang K, Wang H, Yang LT, Jin Q (2020) Deep learning enhanced human activity recognition for internet of healthcare things. IEEE Internet Things J
7. Alleghany M, Baker T, Al-Jumeily D, Hussain A, Mustafina J, Aljaaf AJ (2020) Prospects of the machine and deep learning in analysis of vital signs for the improvement of healthcare services. In: Nature-inspired computation in data mining and machine learning. Springer, Cham, pp 113–136
8. Sendak M, Elish MC, Gao M, Futoma J, Ratliff W, Nichols M, Bedoya A, Balu S, O'Brien C (2020) "The human body is a black box" supporting clinical decision-making with deep learning. In: Proceedings of the 2020 conference on fairness, accountability, and transparency, pp 99–109
9. Ali F, El-Sappagh S, Riazul Islam SM, Kwak D, Ali A, Imran M, Kwak K-S (2020) A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Inf Fusion
10. Naz H, Ahuja S (2020) Deep learning approach for diabetes prediction using PIMA Indian dataset. J Diabetes Metab Disord 19(1):391–403
11. Pitoglou S (2020) Machine learning in healthcare: introduction and real-world application considerations. In: Quality assurance in the era of individualized medicine. IGI Global, pp 92–109
12. Uddin Md Z, Hassan MM, Alsanad A, Savaglio C (2020) A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare. Inf Fusion 55:105–115
13. Jain V, Chatterjee JM (eds) (2020) Machine learning with health care perspective: machine learning and healthcare, vol 13. Springer Nature, Berlin

Design and Analysis of Metamaterial Waveguide Antenna for Broadband Applications

Smrity Dwivedi

Abstract Present paper proposes an antenna-based metamaterial waveguide for broadband applications which works on 7.8 GHz for its return loss. Analysis has been developed X band waveguide which covers coupled split ring resonator having broadside version and C-band waveguide to excite the metamaterial structure as a waveguide to coaxial transition. This structure is numerically analyzed using commercial software CST studio tool. Structure is simulated on dielectric substrate RT duroid with dielectric constant 2.33. Obtained return loss and VSWR are −35 dB and 1.2, respectively.

Keywords Metamaterial · SRR · Waveguide · Transition · Return loss

1 Introduction

Planar guiding structures have changed the era of design in which the rectangular waveguides used in so many of applications area are replaced by its miniaturization version. Although, waveguides are still being needed in many applications in defense as well as research. Waveguides have basically been providing the advantages of low loss and high power handling capability. Planar antennas like patch antennas have certain disadvantages such as substrate mode losses, spurious radiation from various feeding networks, and poor polarization purity as well as narrow bandwidth.

From electromagnetic theory, it is well known and discussed in many research papers and books that for propagation of waves possible when through a waveguide having half wavelength width [1]. Classical waveguide miniaturization was already done by filling dielectric material in to waveguide structures [2–5], and its size is equal to one by square root of dielectric constant. Hence, air-filled waveguides have a low loss compared to the equivalent dielectric-filled waveguide and also high εr dielectric-filled waveguides do not radiate from the open end as effectively as air-filled waveguides. Since there are so many techniques have been adopted to

S. Dwivedi (✉)
Department of Electronics Engineering, IIT-BHU Varanasi, Varanasi, India
e-mail: sdwivedi.ece@itbhu.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_42

reduce the size of waveguide so that it can be used in several applications due to its huge applications as such as magnetic conductors to make a waveguide walls for application in electromagnetic and it is artificial [6] with transverse electromagnetic mode propagation effectively. This process could also be used for reduction of the waveguide width. Recently, Marques and Hrabar have proposed a little different waveguide for broad application area [7, 8]. Metal rectangular waveguide which is loaded with resonant structures is given by them is used in many applications, and it is periodic structure which is also known as split ring resonators (SRR) [8–13].

Basically, SRR belongs to a broader class of materials which is known electromagnetic metamaterials. Metamaterials is famous for its big interest due to their unique responses to electromagnetic radiations, which are generally not encountered in their natural entity. These materials are engineered by embedding precise metallic inclusions or inhomogeneities in given structure like substrate at microwave frequencies. The special characteristics of MTM is originates their properties from their structure and not inherit from chemical composition of their constituents. Metamaterials are fully synthetic materials, and it is developed to get unique properties not normally found in nature. It was artificial dielectric in terms of electromagnetics. Appropriately designed metamaterials can affect the electromagnetic radiation, which is not found in bulk materials. Basically, metamaterials are divided into two major categories, resonant and non-resonant types, depending upon the oscillation of waves as well as periodicity of structures and reversal of Snell's law and amplification of evanescent waves that is not possible normally that can be realized. Hence, since such materials are not readily available in nature usually, and Veselago's predictions did not become popular in research community until Caloz [11] had succeeded in constructing a material using SRR and straight conducting wires which worked for negative permittivity and permeability. Due to that, it covers lot of number of frequency band because of not having positive refractive index.

Analytical as well as mathematical proof of wave propagation in circular and rectangular waveguides and the concept of cutoff frequency were initially given by Belov [14]. Before that, Heaviside had proposed wave propagation through closed hollow tubes like vacuum tubes, etc., but idea was rejected just because wires were required for wave propagation [15] always and it is mandatory. Experimental calculation of the rectangular and circular waveguides was done in by George C. Southworth and W. L. Barrow, respectively. Waveguides have been used as guiding structures in antenna like horn antennas, radar, and in many microwave applications for over several decades.

2 Analysis

In today's era of handheld and compact, lightweight integrated micro and nanodevices, integratable planar guiding structures such as slot lines, micro strip, fin line, and coplanar lines have replaced the heavy bulky waveguides. But in many certain applications, like feeding networks for large antenna arrays, more than

dual frequency antennas, specified radar applications, satellite communications, and broadcasting, where the use of waveguide is almost necessary and unavoidable. All present waveguide offers many advantages over the spurious radiation and lossy coaxial or microstrip lines, and waveguides are completely shielded without any radiation which has low loss and high power handling capability. Planar antennas like patch antennas have certain disadvantages like substrate mode losses, spurious radiation from feeding networks, narrow bandwidth, and poor polarization purity. Apart from that, the three-dimensional nature of waveguide antennas, which is more prominent although, it is difficult to integrate this antenna with planar circuitry. Manufacturing costs of the waveguides are high usually because of the material and its volume. The biggest disadvantage of waveguide is its size and weight at low microwave frequencies that is creates a difficulty to get compact sizes. Apart from these disadvantages and due to increase demand for a compact, lightweight and miniaturized waveguide which can work at low microwave frequencies, that became possible after the evolution of electromagnetic metamaterials which came into research area, and it has applications supporting backward waves in guiding structures were studied by a number of researchers [6–8, 16–18].

To design this structure, first design parameters are calculated which is given in Table 1.

Using CST microwave studio, design has been drawn for the structure mentioned. Figure 1a represents the front view and b represents the back view of structure where BC-SRR is clearly visible. In this design, the C-band waveguide is coupled as transition structure and other side is air box to excite the structure.

This is very unique featured material to design SRR due to its negative permittivity and negative permeability. Metamaterial is an engineered material having with periodic structure gives low loss and high gain as well as light weight material. It also gives the high bandwidth and good for getting optical behavior. Figure 2 shows the negative permittivity and negative permeability of metamaterial to design the SRR.

Figure 3 represents the surface current in designed structure. From the plot, it is clear that current is dense toward the transition waveguide which is having few SRR inside it.

Table 1 Parameter and their values used in design

Sr. No.	Parameter	Value
1.	Radius of the BC-SRR	2 and 3 mm
2.	X band waveguide box	35 mm × 25 mm × 15 mm
3.	C-band transition waveguide	22.86 mm × 60 mm × 10.16 mm
4.	Height of substrate	2.25 mm
5.	RT duriod	2.33

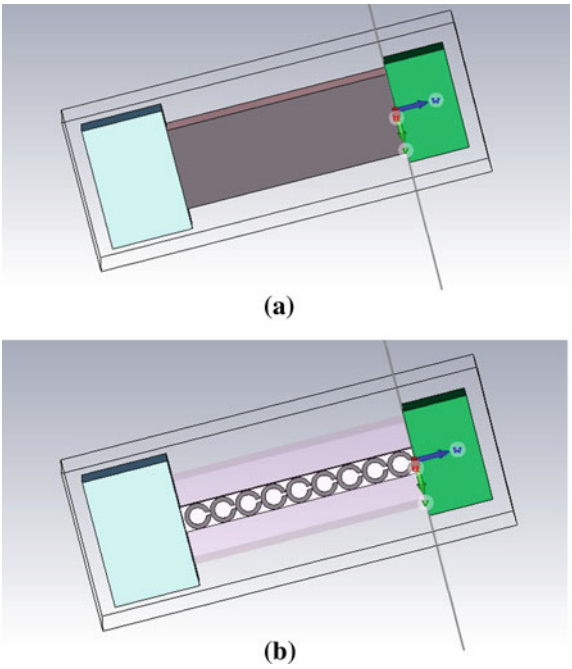


Fig. 1 Schematic view of metamaterial waveguide antenna, **a** front view and **b** back view

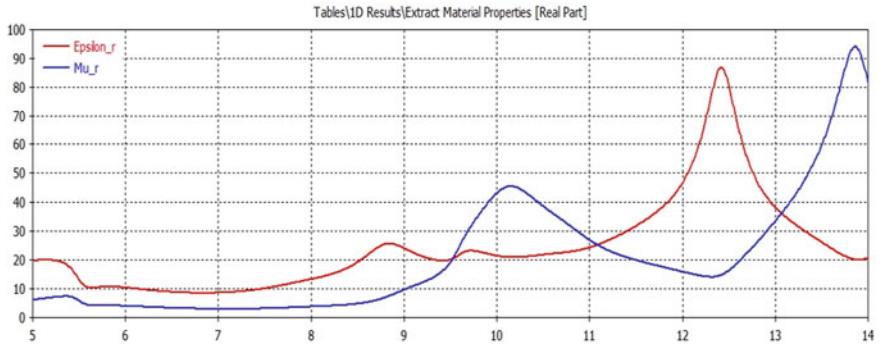


Fig. 2 Epsilon and mu plots for metamaterial structure

From Fig. 4a and b, it is clear that the maximum gain from the given antenna ~10 dBi. From the plot, the side lobes are observed. This is happened because of coupling issue between two waveguides and impedance mismatching.

Figure 5 represent the transmission of power from port to the C-band waveguide.

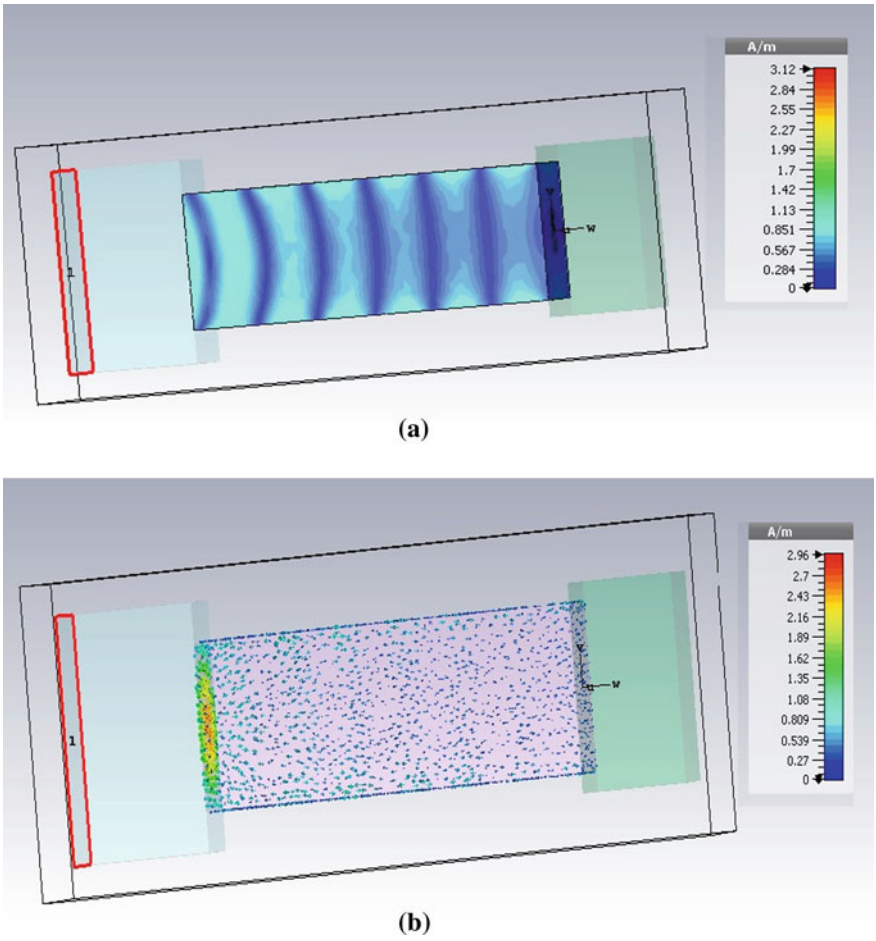


Fig. 3 Surface current in designed structure, **a** contour plot and **b** vector plot

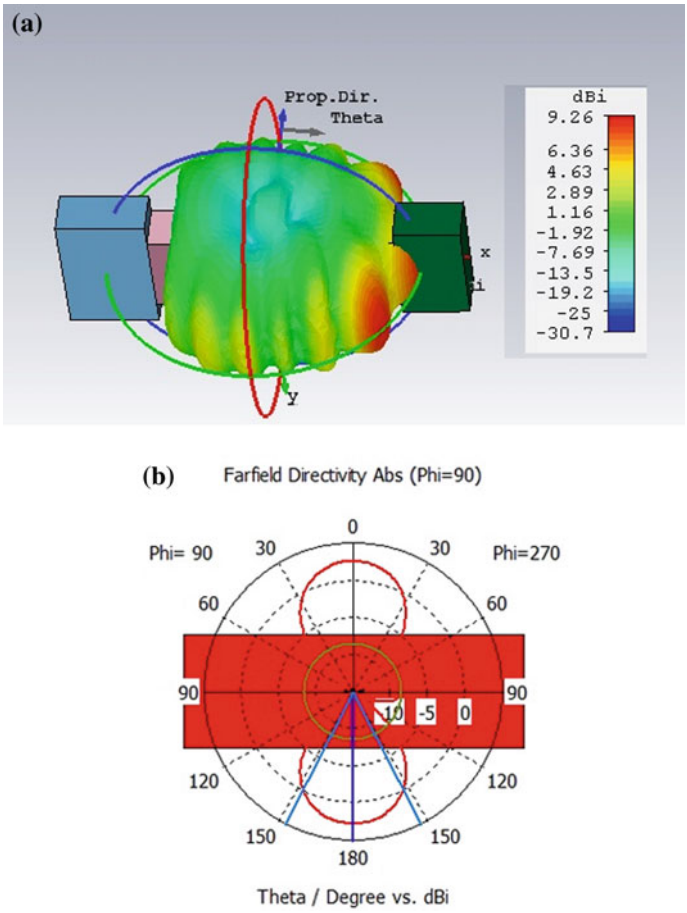


Fig. 4 **a** Radiation pattern 3D plot and **b** radiation pattern in 2D (polar) plot

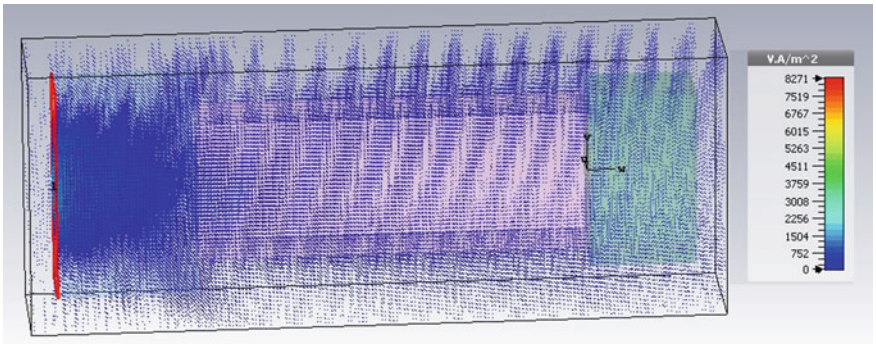


Fig. 5 Power flow from excitation to transition waveguide

3 Conclusion

For broadband application, design has been done. Although it is complicated design but output will be achieved maximum. The return loss is −35 dB and the VSWR is 1.2. This can be improved by improving the coupling between the two X band and C-band waveguide. The gain for this designed structure is ~10 dBi.

References

1. Marcuvitz N (1986) Waveguide handbook. IET
2. Konishi Y (1991) Novel dielectric waveguide components-microwave applications of new ceramic materials. Proc IEEE 79(6)
3. Chang C-Y, Hsu W-C (2002) Novel planar, square-shaped, dielectric-waveguide, single- and dual-mode filters. IEEE Trans Microw Theory Tech 50(11)
4. Fei-Ran Y, Kuang-Ping M, Yongxi Q, Itoh TA (1999) Novel TEM waveguide using uniplanar compact photonic-bandgap (UC-PBG) structure. IEEE Trans Microw Theory Tech 47(11):2092–2098
5. Shelby RA, Smith DR, Nemat-Nasser SC, Schultz S (2001) Microwave transmission through a two-dimensional, isotropic, left-handed metamaterial. Appl Phys Lett 78(4)
6. Harbar S, Bartolic J, Sipus Z (2005) Waveguide miniaturization using uniaxial negative permeability metamaterial. IEEE Trans Antennas Propag 53(1)
7. Marques R, Martel J, Mesa F, Medina F (2002) Left-handed-media simulation and transmission of EM waves in subwavelength split-ring-resonator-loaded metallic waveguides. Phys Rev Lett 183901–183904
8. Hrabar S, Bartolic J (2003) Backward-wave waveguide based on uniaxial anisotropic negative permeability meta-material. In: Proceedings of ICECOM 2003
9. Smith DR. Electromagnetic metamaterials: towards higher frequencies
10. Smith DR, Padilla WJ, Vier DC, Nemat-Nasser SC, Schultz S (2000) A composite medium with simultaneously negative permeability and permittivity. Phys Rev Lett 84(18):4184–4187
11. Caloz C, Sanada A, Itoh T (2003) Microwave circuits based on negative refractive index material structures. In: Proceedings of EuMc 2003, Munich, pp 389–392
12. Rayleigh L (1903) On the passage of electric waves through tubes. Philos Mag 43
13. Heaviside O (1893) Electromagnetic theory, vol 1
14. Belov PA (2005) Subwavelength metallic waveguides loaded by uniaxial resonant scatterers. Phys Rev E 72:036618. © 2005 by the American Physical Society, 2005
15. Hrabar S, Jankovic G (2006) Scanning leaky-wave antenna based on a waveguide filled with plasma-like ENG metamaterial. In: IEEE MELECON, 16–19 May, Benalmádena (Málaga), Spain
16. Lubkowski G, Damm C, Bandlow B, Schuhmann R, Schußler M, Weiland T. Waveguide miniaturization using spiral resonators and dipole arrays. In: Proceedings of the 36th European microwave conference
17. Hrabar S, Eres Z, Bartolic J. Capacitively loaded loop as basic element of permeability meta-material
18. Eleftheriades GV, Antoniades MA, Grbic A, Iyer A, Islam R. Electromagnetic applications of negative-refractive-index transmission-line metamaterials

Secure Messaging Application with Live Translation and Privacy

Mrunal Dhomane and Ashlesha Nagdive

Abstract We all tend to use the native language as much as possible. Even though English is the most popular and professional language, merely 20% of the population can communicate using it. In this planned project, We have developed a secure messaging application with a translation feature considering privacy issues of the user. The goal of this application is to provide messaging functionality with a translation feature (translate the message from English to Hindi and vice versa), maintain privacy by blocking screenshot/s and screen recording, encrypting and decrypting message/s. To implement this, I have used Android OS as it is used by nearly 70% of the world's population. We have also used the firebase platform and Google ML kit. Google ML kit helps to achieve translation features.

Keywords Text translation · Encryption · Decryption · Privacy

1 Introduction

Instantaneous messaging app is a conduit for our personal information. It now has global users exchanging their personal information. Instant messaging applications have great features and facilities which make them well-liked among smartphones users. WhatsApp can be used as an instant messaging application but the limitation is that we cannot translate the text in the app and cannot block the screenshots or screen recording for our privacy concerns. For translation, we have Google translate but we have to manually feed the text. This project confounds all of these. I have integrated different platforms and services to form a solution for the above drawbacks. Android provides an open-source application environment and platform for the developers.

The Android core operating system is based on the Linux kernel which ensures security. "Tattle" is the application that is presented in this work that provides the

M. Dhomane (✉) · A. Nagdive
G H Raisoni College of Engineering, Nagpur 440016, India
e-mail: mrunal.dhomane.it@ghrce.raisoni.net

A. Nagdive
e-mail: ashlesha.nagdive@raisoni.net

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_43

features which are not present in WhatsApp. This application ensures the full encryption (AES-128 bit) to all the text messages between two users, blocks the screenshots and screen recording, and translates the message in the app itself without using Google Translate.

2 Literature Review

There are many chatting applications available on the Play Store. Recently, applications have started highlighting "Secure chat application" and "User-friendly application". These applications profess by the owners/organization that they have privacy concerns and are actually user friendly. Unfortunately, there are some points where these applications fail to meet the requirements. To comprehend different messaging applications few applications have been picked based on their Play Store rating.

2.1 WhatsApp

It is one of the most popular and preferred messaging application present in the market today. With end-to-end encryption, it provides the most acceptable security at the current point and has over 2000 million monthly active users [1]. It is present on both the platforms, i.e. Android and iOS. Numerous tasks can be done by using this single application. Sharing multimedia, payments, Web application, peer-to-peer messaging, and group messaging [2]. It doesn't provide an in-app translation feature for native users and allows taking screenshots and screen recordings which may be considered a privacy breach for some user.

2.2 Snapchat

Snapchat is an American multimedia instant messaging application developed by snap Inc. Snapchat was initially focused on private and person-to-person photo sharing. Also, there can be a transfer of text messages, voice messages, photos, and videos between two users or groups. It notifies the user if someone has taken a screenshot of the chat but doesn't block it. According to Statista as of October 2021, Snapchat has over 538 million monthly active users [1]. Chats are deleted by default once both the users leave the Chat or can change to "delete after 24 h" for deleting the chat after 24 h.

2.3 Secure Messaging Application

A paper on end-to-end secure messaging service by Somen Nayak and 17 others has proposed a protocol that uses a user-defined password for SHA-2 hash generation. SHA-2 hash generation is used as a key of AES-256 encryption to encrypt the message during transmission of the message. They have designed HTTP server for holding user data and controlling authentication. The security part is implemented using AES-256 module [3].

2.4 Google Translate

The Google Translate app can translate many languages either through text or voice. It even has an OCR feature that allows user to point their smartphones at the text or sign which is not written in their native language. Along with English, a small sampling of the many languages is included in the app. For example, German, Spanish, Greek, Chinese, Japanese, Korean, Latin, etc. It also has a real-time conversations feature that helps users to talk with someone, even if the languages are distinct [4]. An article on CNET states that "Google Translate provides a billion translations a day for 200 million users, the company revealed here Friday at its Google I/O show for developers" [5]. The only unfavorable point is, we have to manually feed text in it to translate.

2.5 Android Live Text Recognition and Translation Application Using Tesseract

Dr. S. Revathy and Sakhayadeep Nath offered a translation application using tesseract leptonica for OCR and Google-API-translate to translate the text into the desired language. It is developed on UNIX base OS platform; however; the paper states "The accuracy in recognizing handwritten text is low and needs additional work." It is the practical implementation of the related work by Harneet Singh and Anmol Scahan [6].

2.6 Mobile Language Translator

A paper proposed by Sim Liew Fong and 3 others on language translation (English to Bahasa and vice versa) for helping tourists with hassle-free communication. The proposed work intended to overcome the language barrier among tourists and help them getting correct information [7].

3 Methodology

With the help of the latest technologies and services available for Android development, the proposed work can be implemented in Java. Firebase platform and the android studio have been used in the process of implementing and developing the application. Firebase platform provides database, authentication service, and storage for the growth of the project. Google ML kit provides text translation with four key capabilities:

- Broad language support
- Proven translation models
- Dynamic model management
- Runs on the device itself.

The application has a very simple working structure. It is based on model view controller (MVC). Model view controller is a software design pattern commonly used for developing user interfaces that divide the related program logic into three interconnected elements (Fig. 1).

- **Model**—Contains only the pure application data.
- **View**—Presents the model's data to the user.
- **Controller**—Exists between view and model. It listens to event triggered by the view, and the reaction is to call a method on model.

The application uses AES encryption to encrypt the message to provide privacy to the users. A paper on the implementation of advanced encryption standard algorithm states that it should support either 128 or 192 or 256 bits [8] (Fig. 2).

The application uses 128-bit encryption to encrypt the message before uploading it to the database. Once the plain text is converted to ciphertext, the app uploads the ciphered text to the database (Fig. 3).

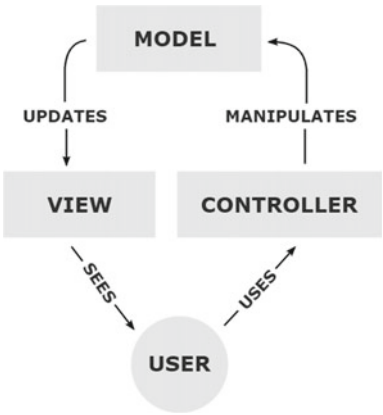


Fig. 1 Model view controller

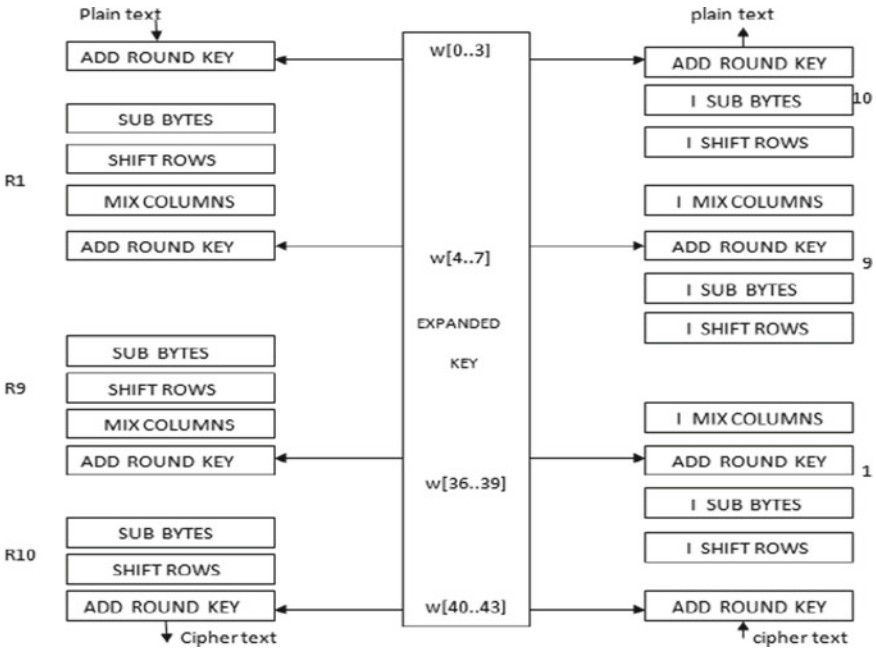


Fig. 2 General structure of AES algorithm

Fig. 3 Encrypted message in the database

The receiver fetches the encrypted text from the database, the application decrypts the message, and the user can read the message (refer Fig. 4). On the completion, the users can translate the message in their native language using Google ML kit. As stated above, one of the key capabilities of ML kit is proven translation models. "The models are powered by same models used by the Google Translate app's offline mode" [9]. The kit is based on Google neural machine translation (GNMT) that uses

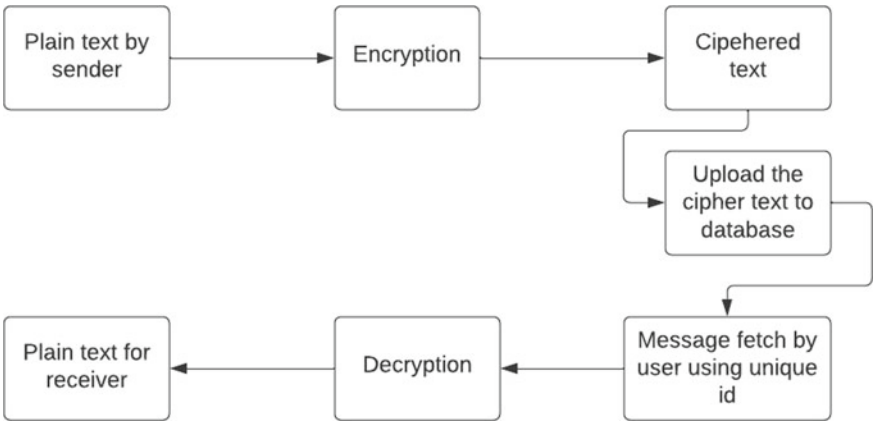


Fig. 4 General working structure of the application

an artificial neural network to increase fluency and accuracy in Google Translate [10]. The proposed application can translate (English to Hindi and vice versa) without sending the data to the server or any other application that prevents privacy breaches.

The application can block screen recording and screen shots for privacy concerns of the user. The user can permit to take the screenshot by updating the status in the app settings.

4 Results

The features proposed in the paper are shown below figures. These outputs will help in better understanding of the flow and working structure. The main features are text translation and blocking screenshot.

4.1 Text Translation

Long press on the message box translates the text from English to Hindi using the Google ML kit. Both sender and receiver can translate their message using this feature. This feature can work offline once the conversation is loaded and doesn't make changes to messages on the database. Refer figures mentioned below (Figs. 5 and 6).

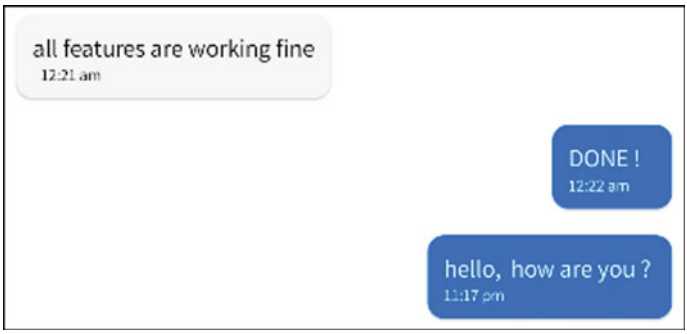


Fig. 5 Original message

Fig. 6 Translated message

4.2 Screenshot Blocking Feature

Users have full discretion over this attribute. The user can update their preference from app settings. Once the preference is updated on the database, the app gets the preference in the background, and depending upon the result, it allows or denies the screenshot or screen recording from the receiver end which takes care of the privacy of the user. App gives the toast stating "Can't take screenshot due to security policy" if the permission is denied (Fig. 7).

These two are the supreme goal of the application which have been successfully implemented. There is no limitation to the number of messages and translation.

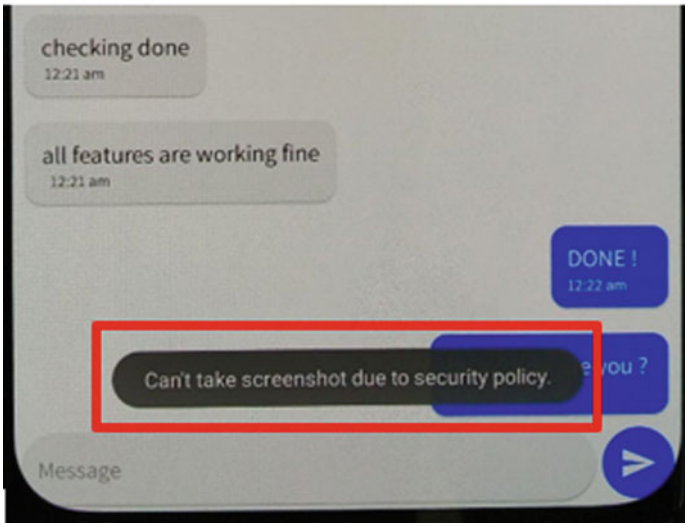


Fig. 7 Screenshot permission denied

5 Future Scope

Any technical innovation adds value to itself with various expansions and advancements to the existing ideas over time. This project also comes as an origin to various value additions that could make a tremendous scope of this technology in the future.

Though this technology is initiated with the translation facility to limited languages at present, the concept of multilanguage support can be expanded in the future where messages can be translated into 50+ languages to expand the range of services for the users. There may be a user who wants to use it in a different language. Therefore, the reach of this technology shall act as a boon for users across the world by supporting 50+ languages through this technology.

On top of that text-to-speech (TTS) is a popular feature that helps people who have difficulty in reading the text on the screen. The automated voice of the TTS shall read the text aloud so that the user can hear the message conveyed and can also read it. This shall act as a Tech Friendly source for people with a challenge in reading the text. One of the TTS tools named optical character recognition which can read the text aloud from images can also act a vast scope in the future. The aforesaid features can be added in the future with proper resources and with in-depth knowledge about the technology.

6 Conclusion

The intensive efforts toward the creation of this idea shall lead to the application acting as a user-friendly device that can successfully translate the text from English to Hindi and vice versa. This idea also evolves around the privacy of the users with a security thus encrypting the message before uploading on the database. It blocks screenshots and screen recordings. We acknowledge that this application will be useful for a more user-friendly environment and in many situations. To achieve the functionality, application uses many online services and open-source libraries.

In the future, new improvements and functionality can be implemented to provide more features like access multiple language support, more security and privacy to the users using the latest algorithms and privacy techniques. Using this application, we can grow the android environment among people so we can maximize user support and minimize their problems to improve development.

References

1. Statistic of most popular application/s. https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/
2. Whatsapp features. https://www.whatsapp.com/features
3. Nayak S, Das S (2017) An application for end to end secure messaging service on Android supported device. In: 2017 8th IEEE Annual information technology, electronics and mobile communication conference (IEMCON), 23 Nov 2017. IEEE
4. Google translate. https://in.pcmag.com/reviews-and-price-comparisons-from-pc-magazine-of/134966/dont-speak-the-language-how-to-use-google-translate
5. Google translate article on CNET. https://www.cnet.com/tech/services-and-software/google-translate-now-serves-200-million-people-daily/
6. Revathy S, Nath S (2020) Android live text recognition and translation application using tesseract. In: 2020 4th International conference on intelligent computing and control system (ICICCS), 13 June 2020. IEEE
7. Fong SL, Elfaki AO (2012) Mobile language translator. In: 2011 Malaysian conference in software engineering, 26 Jan 2012. IEEE
8. Pitchaiah M (2013) Implementation of advanced encryption standard algorithm. Int J Sci Eng Res (IJSER) 3
9. Google ml kit. https://developers.google.com/ml-kit/language/translation
10. Google translate. https://en.wikipedia.org/wiki/Google_Neural_Machine_Translation

A Survey of Different Approaches for Word Sense Disambiguation

Rasika Ransing and Archana Gulati

Abstract Analysis of textual data helps to understand the perception of people by studying the various senses of words in the text. The sense of a polysemous word varies as per the context in the sentence. The technique for determining the correct interpretation of a polysemous word according to context is called as Word Sense Disambiguation (WSD). Recently, researchers have proposed many algorithms to solve this linguistic ambiguity problem in different languages. In this paper, we give a general summary of current trends in WSD in terms of automation of disambiguation approaches. We also mention the challenges and future directions for research for WSD systems. We also propose a system based on these future directions for research, which may increase the accuracy of WSD system for Indian languages.

Keywords Supervised learning · Word Sense Disambiguation (WSD) · Unsupervised learning · Knowledge-based · Machine Learning · Semi-supervised learning · WordNet · Natural Language Processing (NLP)

1 Introduction

Natural languages are filled with ambiguity. Various types of ambiguity are classified according to their nature. It is worth addressing a few sense-related ambiguities, such as Polysemy and Homonymy in the context of Word Sense Disambiguation (WSD) [1].

– Homonymy is a term that refers to a set of words that have the same spelling and sound but distinct meanings. These senses may be completely unconnected to one another. For example, the term “left" can refer to the “past tense of the verb depart" or the “opposite of right."

R. Ransing (✉)
Datta Meghe College of Engineering, Navi Mumbai, Maharashtra, India
e-mail: rasikaransing275@gmail.com

A. Gulati
School of Business Management, NMIMS University, Mumbai, Maharashtra, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_44

– Polysemy is a type of homonymy in which many meanings of the same word are connected to one another. These senses have a common ancestor. For example, the term “mouth” can refer to both “the mouth of an animal” and “the mouth of a cave or river”. Because it is difficult to discern between closely related senses, polysemy is more challenging for WSD to manage than other homonyms.

Conceptually, Word Sense Disambiguation (WSD) is the process of determining the exact interpretation of a polysemous term in a given context [2]. WSD consists of two steps - the count of senses of ambiguous words is determined in the first phase using sense inventory, and the proper sense of the ambiguous word is assigned in the second step [3]. Many Natural Language Processing (NLP) applications employ WSD, either directly or indirectly. Sentiment Analysis, Machine Translation, Information Retrieval, Text summarization, Text Entailment, Semantic Role Labeling are some of the applications in NLP where WSD is employed. WSD is often referred to as the “heart" of NLP. Here are some of the approaches of resolving word sense ambiguity based on a variety of training materials:

– Corpus-based approach: The corpus-based method uses feature vectors to express linguistic information for each sentence’s context (e.g., the use of an ambiguous term). These characteristics might be of a different type: collocations of words, part-of-speech labels, keywords, domain and topic information, grammatical connections, and so on. There exist three possible strategies to corpus-based WSD systems:
  – Supervised approaches: It uses machine learning techniques trained on sense-tagged corpora.
  – Unsupervised approaches: It does not necessitate the use of labeled corpora. It does not classify the instances but clusters the instances.
  – Semi-supervised approaches: It employs a small amount of sense-tagged corpora while also employing a large amount of unlabeled corpora.
– Knowledge-based approaches: To interpret the sense of the target word, it uses a large number of lexical resources like ontologies, machine-readable dictionaries, and thesaurus.

2 Disambiguation of Word Senses: Different Approaches

We describe the different approaches for WSD based on whether training data is required or not, in this section.

2.1 Supervised Approaches

Navigli and Pasini [4] proposed Train-O-Matic which uses knowledge-based language-independent approach. The system consists of 3 steps: (i) Lexical pro-

filing where the system computes a Personalized PageRank vector for every vertex (i.e., synset) in the semantic network that gives its lexical-semantic profile, (ii) Sentence scoring where the system computes a distribution over its senses given a target term and a group of sentences in which it is present, and (iii) Sentence ranking and selection where according to the gap between the first and second senses in the distribution generated in the prior phase, Train-O-Matic rates every phrase for a sense of a word.

AMuSE-WSD [5] as proposed by Orlando et al., performs sense disambiguation in 40 languages through neural network. The AMuSE-WSD model, when given a word in context, (i) builds a contextualized representation of the word as the average of the hidden states of the last four layers of a pre-trained Transformer encoder, (ii) applies a non-linear transformation to obtain a sense-specific hidden representation, and eventually (iii) calculates the distribution of output score over all the possible senses of a sense inventory. Their results show improvement as compared to other WSD systems that are based on non-neural machine learning or graph-based heuristics algorithms.

It Makes Sense (IMS) [6] is a supervised English all-words WSD system proposed by Zhong and Tou Ng. The versatile framework of IMS enables users to combine various pre-processing techniques, extra features, and classifiers. The classifier is a linear support vector machine with several knowledge-based characteristics. Any text can be entered into the system. IMS disambiguates the sense of each content word (noun, verb, adjective, or adverb) in the input text and outputs a list of the senses of word, with each sense s_i assigned a probability based on the likelihood of s_i existing in that context.

The supervised Word Sense Disambiguation API, SUPWSD [7] written in Java is proposed by Papandrea et al. (WSD). The implementation of a supervised WSD system, as well as a NLP framework for feature extraction and pre-processing, are included in this toolkit. SUPWSD is made up of four primary modules that are used in both training and testing phases—Input parsing, Text pre-processing, Features extraction and Classification.

In Ajai Kumar et al. [8] generate a sense-tagged corpus. Five feature sets are constructed using this corpus, and all three classifiers are trained on them. They discovered that Support Vector Machine-Based Classifier did well on the corpus, followed by Neural Network-Based Classifier and Naive Bayes-Based Classifier. They came to the conclusion that a system for Hindi word sense disambiguation can be created utilizing the feature set that produces the best results across all classifiers.

For supervised word sense disambiguation, Yamaki et al. in [9] provides an approach that uses sentence similarities from context word embeddings. An N-dimensional vector with N similarities between each pair of sample sentences is added to a basic feature vector if there are N example sentences in the training data. This new feature vector is utilized to train and identify a classifier.

The gloss selection target for supervised WSD was proposed by Yap et al. in [10], who defined WSD as a problem of ranking of relevance based on context-gloss pairs. They also propose fine-tuning BERT to pick the most likely sense description

given a context phrase and a list of candidate sense definitions using a sequence-pair ranking task. They also present a data enhancement strategy for WSD based on existing WordNet example sentences.

2.2 Unsupervised Approaches

In Gogoi et al. [11] have proposed applying a genetic algorithm to all possible senses of a given ambiguous word taken from WordNet for Assamese WSD. Wu-Palmer's similarity measure was proposed to be used to all extracted senses for the fitness function, and disambiguation was done by choosing the sense with the highest similarity score as the winner.

A huge WSD dataset is presented by Heo et al. in [12], which is automatically created from the Oxford Dictionary. A WSD model is also offered, which derives the meaning of every word based on its part-of-speech in the context. Additionally, to achieve a satisfactory performance, a hybrid sense prediction approach is presented that categorizes the less often utilized senses independently.

In Song et al. [13] proposes to enhance the sense representations which enables pre-trained language models to distinguish between the word senses in better way by enhancing word sense clustering for each word type. Enhancing the sense representations is done by including example phrases, synonyms or sentences showing use of word senses, and sense gloss of hypernyms. The authors exhibit that including extra information enhances the performance of WSD. The authors propose a binary sentence pair classification model that is built upon RoBERTa with focus on sense representation embellishment.

In [14], it was proposed to use word embedding to discover a solution for WSD for Hindi literature by Archana K. and D.K. Lobiyal. This exercise has two parts: creating a word embedding and using cosine similarity to find an suitable meaning for the term. The word embedding in this system is developed using two popular architectures of word2vec: Continuous Bag-Of-Words and Skip-Gram.

For Punjabi language WSD system, two methods for deep learning - long short-term memory (LSTM) and multilayer perceptron (MLP) were independently evaluated on the word vectors of 66 ambiguous Punjabi nouns by Parteek Kumar in [15]. The inputs to the deep learning techniques were the word vectors that were created from a corpus of Punjabi language that had been manually sense-tagged.

2.3 Semi-supervised Approaches

In Duarte et al. [16] evaluates semi-supervised WSD by taking into consideration various graph-based Semi-Supervised Learning (SSL) algorithms with features created by word embeddings from FastText, Word2Vec, GloVe, ELECTRA, and BERT models merged with word context and parts-of-speech tags. To perform disambiguation

of classical lexical sample WSD datasets, the authors examined various combinations of models for word embedding, measures of similarity for graph building, and Semi-Supervised Learning classification algorithms. The findings of this experiment show that SSL algorithms performed similarly to supervised learning algorithms, and that the ELECTRA models outperformed other SSL embeddings.

In Sousa et al. [17] aim to adapt semi-supervised algorithms for WSD using Word2Vec, BERT, and FastText word embeddings as input, as well as part-of-speech tagging. The authors perform a review of four SSL models based on graph to examine the impact of their hyperparameters on the outcomes, and the distances used to create the graphs, the percentages of labeled data, and the architectural changes in word embeddings.

In Yuan et al. [18] propose the use of LSTM, a sequence learning neural net, to better capture the text's syntactic and sequential patterns. LSTM is used in a semi-supervised label propagation classifier to compensate for the absence of training data in all-words WSD. The authors propose two WSD algorithms that merge (1) A huge unlabeled text corpus used to train LSTM neural network language models, (2) data in the form of example phrases that has been labeled, and (3) unlabeled data in the form of additional sentences.

In a decision rule algorithm, Garigliotti [19] emphasizes important parameters such as the initial labeling of examples, the formalization of rule confidence, and the criteria for approving a decision rule. The author then offers a weakly supervised version of the method, as well as a pseudo-word-based strategy for assessing the impact of these variables.

To make efficient use of tagged and unlabeled samples, Yatabe et al. in [20] suggested a semi-supervised WSD method based on graph embeddings. For WSD, the graph structure between examples is efficient, and using a graph structure derived by fine-tuning BERT in the suggested technique is effective.

Scarlini et al. in [21] proposes the ARES (context-AwaRe Embeddings of Senses) system, which is a SSL approach for generating sense embeddings for lexical meanings within a lexical knowledge base that are comparable to contextualized word vectors. By integrating the representational power of a pre-trained language model with the information within a knowledge base, ARES compensates for the lack of manually annotated instances for a major fraction of words' meanings. This allows for the creation of trustworthy representations for those senses that do not appear in manually curated resources, hence enriching the vectors for all other ideas.

2.4 Knowledge-Based Approaches

In Wang et al. [22] assess the contribution of both the word-level and sense-level global context to the disambiguation of an ambiguous word. The authors have observed that the similarity-based system's performance on WSD can be enhanced by relatively large margins by using Context-Oriented Embedding (COE).

Scarlini et al. in [23], propose the SENSEMBERT system, which is a knowledge-based method for creating sense embeddings in multiple languages. To build high-quality latent semantic representations of word senses in different languages, this method combines language modeling and the large quantity of knowledge included in a semantic network. In this system, the authors use lexical-semantic information from a knowledge base, such as BabelNet, and an encyclopedic resource, such as Wikipedia, instead of manually tagging corpora. Despite the fact that SENSEMBERT does not use annotated data, the authors show that it achieves state-of-the-art performance on multilingual WSD jobs and is competitive with the best supervised models on English.

In Luan et al. [24] suggest a technique to leverage machine translation to enhance the performance of a basic WSD system. This method is language agnostic and produces cutting-edge outcomes in both English all-words and multilingual WSD. The authors offer BABALIGN, an accurate bitext alignment technique guided by multilingual lexical correspondences from BabelNet, to enable the usage of lexical translation information. In order to bridge the gap between the performance of knowledge-based and supervised WSD the syntagmatic information is exploited.

SyntagRank is a knowledge-based WSD system proposed by Scozzafava et al. in [25], which is accessible via a RESTful API and a Web interface. SyntagRank uses disambiguated pairs of cooccurring terms from SyntagNet [26], a lexical-semantic combination resource, to perform cutting-edge knowledge-based WSD in a multilingual scenario. SyntagRank uses the Personalized PageRank (PPR) algorithm to identify the most appropriate sense of a given word in context. The authors work toward achieving state-of-the-art knowledge-based performance across five languages: English, German, French, Spanish, and Italian.

In Wang et al. [27] propose a comprehensive knowledge-based WSD framework which models the problem of hidden semantic space and semantic path behind a text. This method is based on the WordNet Knowledge Base (KB), which uses PageRank and Latent Semantic Analysis (LSA) to model the semantic path and semantic space, respectively. WSD is accomplished in this system with the help of knowledge graph and information retrieval in order to obtain better knowledge exploitation.

In Kwon et al. [28] proposes a knowledge-based system for word sense disambiguation with an aim to use word similarity to identify a better technique to remove irrelevant data. In this system, two methods are adopted by the authors. First, the representation of word vector is utilized to identify word similarity, and the encoding approach takes into account the graphical semantic links from the lexical knowledge bases. Second, an effective technique for retrieving contextual terms from a text in order to analyze an uncertain word using word similarity.

In Chaplot et al. [29] proposes a WSD system that scales linearly with the count of words in the context by leveraging the formalism of topic model. In order to disambiguate a word this system utilizes the whole document as the context instead of just the current sentence. The proposed approach for this system is a variation of Latent Dirichlet Allocation in which topic proportions are substituted by synset

proportions for a document. The data in the WordNet is utilized by applying a logistic-normal prior to document distribution over synsets and a non-uniform prior to synset distribution over words.

2.5 Hybrid Approaches

In Bevilacqua and Navigli [30] propose a hybrid knowledge-based and supervised approach to WSD, named Enhanced WSD Integrating Synset Embeddings and Relations (EWISER). It is a neural supervised architecture. It can leverage knowledge by embedding data from the WordNet Lexical Knowledge Base (LKB) graph into the neural architecture, as well as employ pre-trained synset embeddings to forecast synsets that are not present in the training set.

After acquiring sense embeddings from enriched WordNet glosses, Ming Wang and Yinglin Wang in [31] propose the Synset Relation-Enhanced Framework (SREF), which employs sense relations to improve sense embedding and a try-again technique to reimplement WSD. In this framework,the authors, create a gloss enhancement for those synsets having a brief definition in WordNet so that each synset can acquire a trustworthy sense embedding using BERT characteristics. The contribution of distinct synset relations in WordNet is then investigated using these embeddings to learn relation-enhanced sense embeddings. After running a closest neighbor technique against context embedding of an ambiguous word and relation-enhanced embeddings of potential senses of the words in the first WSD, a try-again process is applied to the top two contending senses utilizing synset relations and the super-sense category.

Yin et al. in [32], provide a framework for word sense disambiguation that includes knowledge-based and supervised learning techniques. This framework makes use of knowledge to create sense-labeled data for supervised learning automatically. Initially, WordNet is used to build a semantic knowledge base named DisDict, which gives improved feature terms that emphasizes the contrasts between word senses, or synsets. DisDict then automatically generates fresh sense-labeled data from unlabeled data. The data that is generated, along with manually labeled and unlabeled data, is then given to a framework of neural network that does both unsupervised and supervised learning to model the semantic relationships between feature words, synsets, and their contexts.

3 Challenges in Word Sense Disambiguation

Many researchers have explored various techniques for Word Sense Disambiguation in different Indian as well as foreign languages other than English. Following are the challenges and open questions currently available for research with respect to Indian languages [2, 8, 11, 14, 15, 28, 33–36]:

- Indian languages are highly inflectional due to which the morphological analyzers fail to correctly lemmatize the words.
- There is enormous semantic variety in Indian languages.
- In various Indian languages, some sentences are so lengthy that they contain a huge amount of irrelevant information.
- In various Indian languages, some sentences are very short due to which it is not possible to retrieve sufficient information from them.
- Spelling errors in languages directly affects the capacity of understanding meaning of the sentences.
- Indian languages are found to be less resourced. The Lexical Knowledge Base like WordNet for Indian languages is not a comprehensive reference for all types of semantic information in words. Hence, there is a necessity to build language-specific WordNet which consists of all the possible information regarding the word senses and the semantic relations among the word senses.
- Resources can be enhanced and tools such as POS tagger, Named Entity Recognizer, Lemmatizer, Stemmer, etc. can be improved which will hence increase the accuracy of WSD.
- In case of unsupervised approaches, sentence clustering techniques can be enhanced, since in comparison to a paragraph or document, the context of a sentence is much smaller.
- It is observed that strategies based on lexical similarity do not perform well. Hence, work must be done for semantic relatedness-based strategies.
- In certain cases, the real meaning of a sentence is different from the literal meaning of that sentence, like Sarcastic sentences. Deep learning techniques can be utilized in such cases to disambiguate the senses.
- Manually creating high-quality full-coverage training data is a time-consuming and exhausting procedure. Hence, there is a necessity to develop reliable methods to automate this process.
- Current WSD systems that are being designed have better performance in disambiguating Nouns and Verbs. There is a necessity to design systems that perform well for disambiguation of Adverbs and Adjectives as well by considering the contextual information.
- There is a need to develop WSD systems that can disambiguate senses of technical terms like the ones used in domains of Law, Medicine, Engineering, etc. For this purpose Generative Adversarial Networks (GANs) can be used to enrich the gloss set for WSD and hence reduce the dependency on WordNet as WordNet cannot offer gloss of domain-specific terms (technical terms).

4 Proposed Methodology

Figure 1 shows the Proposed Methodology for disambiguating word senses in Indian languages considering the above-mentioned challenges [2, 6, 7, 12, 33]. The system

Fig. 1 Proposed methodology for WSD of Indian languages

accepts text in Indian language and for every word in the text, it will output the sense which has highest probability of appearing in that context. Marathi WordNet developed by IIT, Bombay will be used as sense inventory.

Pre-processing of the input text includes: Splitting of sentences with respect to sentence boundary, Tokenization of the split sentences, Stop word Removal, Part-of-Speech (POS) tagging of the tokens, Lemmatization.

Feature Extraction includes POS tags of surrounding words and usage of N-gram model. This stage will help to decode the sense of a word by understanding the context of the sentences.

Classification will be combination of Supervised, Unsupervised, and Knowledge-Based approaches. If all the three approaches give the same sense as output for a word, then that sense is selected as the sense for that word. If any two approaches give the same sense and one approach gives different sense for a word, then Majority Voting algorithm will be used. If all the three approaches give different senses as output for a word, then the sense with highest confidence will be selected as the sense for that word.

5 Conclusion

This paper gives a summary of the various approaches for Word Sense Disambiguation for different languages. We have classified the various WSD approaches in multiple categories and then surveyed the methods by other researchers. Sense-annotated corpus can be challenging and tedious to acquire in Supervised Learning techniques, and the process needs to be repeated for languages, new domains, and sense inventories. As a result, semi-supervised learning (SSL) techniques, which integrate a modest quantity of sense-annotated data, have become increasingly popular. Unsupervised Learning approaches do not use any sense-annotated data for WSD and hence researchers are working to develop such systems with greater accuracy. Sense inventories and knowledge bases are enhanced so as to improve the accuracy of Knowledge-based WSD. Hybrid approaches combining multiple methods are also being developed to have WSD systems with better performance. We have also proposed a methodology which may help to improve the accuracy of Word sense diasmbiguation for Indian Languages.

References

1. Chaplot DS, Bhattacharyya P (2014) Literature survey on unsupervised word sense disambiguation. IIT Bombay, May 7, p 39
2. Pal AR, Saha D, Naskar SK, Dash NS (2021) In search of a suitable method for disambiguation of word senses in Bengali. Int J Speech Technol 24(2): 439–454
3. Bevilacqua M, Pasini T, Raganato A, Navigli R (2021) Recent trends in word sense disambiguation: a survey. In: Proceedings of the thirtieth international joint conference on artificial intelligence, IJCAI-21, Inc
4. Pasini Tommaso, Navigli Roberto (2020) Train-o-matic: supervised word sense disambiguation with no (manual) effort. Artif Intelli 279:103215
5. Orlando R, Conia S, Brignone F, Cecconi F, Navigli R (2021) AMuSE-WSD: an all-in-one multilingual system for easy word sense disambiguation. In: Proceedings of the 2021 conference on empirical methods in natural language processing: system demonstrations, pp 298–307
6. Zhong Z, Ng HT (2010) It makes sense: a wide-coverage word sense disambiguation system for free text. In: Proceedings of the ACL 2010 system demonstrations, pp 78–83
7. Papandrea S, Raganato A, Bovi CD (2017) SUPWSD: a flexible toolkit for supervised word sense disambiguation. In: Proceedings of the 2017 conference on empirical methods in natural language processing: system demonstrations, pp 103–108
8. Kumar A, Mathur I, Darbari H, Purohit GN, Joshi N (2016) Implications of supervised learning on word sense disambiguation for Hindi. In: Proceedings of the second international conference on information and communication technology for competitive strategies, pp 1–5
9. Yamaki S, Shinnou H, Komiya K, Sasaki M (2016) Supervised word sense disambiguation with sentences similarities from context word embeddings. In: Proceedings of the 30th Pacific Asia conference on language, information and computation: oral papers, pp 115–121
10. Yap BP, Koh A, Chng ES (2020) Adapting BERT for word sense disambiguation with gloss selection objective and example sentences. arXiv preprint arXiv:2009.11795
11. Gogoi A, Baruah N, Sarma SK (2020) Assamese word sense disambiguation using genetic algorithm. In: Proceedings of the 17th International Conference on Natural Language Processing (ICON), pp 303–307
12. Heo Y, Kang S, Seo J (2020) Hybrid sense classification method for large-scale word sense disambiguation. IEEE Access 8:27247–27256
13. Song Y, Ong XC, Ng HT, Lin Q (2021) Improved word sense disambiguation with enhanced sense representations. In: Findings of the association for computational linguistics: EMNLP 2021, pp 4311–4320
14. Kumari A, Lobiyal DK (2020) Word2vec's distributed word representation for hindi word sense disambiguation. In: International conference on distributed computing and internet technology. Springer, Cham, pp 325–335
15. Kumar P (2020) Word sense disambiguation for Punjabi language using deep learning techniques. Neural Comput Appl 32(8):2963–2973
16. Duarte JM, Sousa S, Milios E, Berton L (2021) Deep analysis of word sense disambiguation via semi-supervised learning and neural word representations. Inf Sci 570: 278–297
17. Sousa S, Milios E, Berton L (2020) Word sense disambiguation: an evaluation study of semi-supervised approaches with word embeddings. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, pp 1–8
18. Yuan D, Richardson J, Doherty R, Evans C, Altendorf E (2016) Semi-supervised word sense disambiguation with neural models. arXiv preprint arXiv:1603.07012
19. Garigliotti D (2019) Semi-supervised learning for word sense disambiguation. arXiv preprint arXiv:1908.09641
20. Yatabe R, Sasaki M (2020) Semi-supervised word sense disambiguation using example similarity graph. In: Proceedings of the graph-based methods for natural language processing (TextGraphs), pp 51–59

21. Scarlini B, Pasini T, Navigli R (2020) With more contexts comes better performance: contextualized sense embeddings for all-round word sense disambiguation. In: Proceedings of the 2020 conference on Empirical Methods in Natural Language Processing (EMNLP), pp 3528–3539
22. Wang M, Zhang J, Wang Y (2021) Enhancing the context representation in similarity-based word sense disambiguation. In: Proceedings of the 2021 conference on empirical methods in natural language processing, pp 8965–8973
23. Bianca S, Tommaso P, Roberto N (2020) Sensembert: context-enhanced sense embeddings for multilingual word sense disambiguation. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, issue 05, pp 8758–8765
24. Luan Y, Hauer B, Mou L, Kondrak G (2020) Improving word sense disambiguation with translations. In: Proceedings of the 2020 conference on Empirical Methods in Natural Language Processing (EMNLP), pp 4055–4065
25. Scozzafava F, Maru M, Brignone F, Torrisi G, Navigli R (2020) Personalized PageRank with syntagmatic information for multilingual word sense disambiguation. In: Proceedings of the 58th annual meeting of the association for computational linguistics: system demonstrations, pp 37–46
26. Maru M, Scozzafava F, Martelli F, Navigli R (2019) SyntagNet: challenging supervised word sense disambiguation with lexical-semantic combinations. In: Proceedings of the 2019 conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp 3534–3540
27. Wang Y, Wang M, Fujita H (2020) Word sense disambiguation: a comprehensive knowledge exploitation framework. Knowl -Based Syst 190:105030
28. Kwon S, Dongsuk O, Ko Y (2021) Word sense disambiguation based on context selection using knowledge-based word similarity. Inf Process Manage 58(4):102551
29. Chaplot DS, Salakhutdinov R (2018) Knowledge-based word sense disambiguation using topic models. In: Proceedings of the AAAI conference on artificial intelligence, vol 32, no 1
30. Bevilacqua M, Navigli R (2020) Breaking through the 80% glass ceiling: raising the state of the art in word sense disambiguation by incorporating knowledge graph information. In: Proceedings of the 58th annual meeting of the association for computational linguistics, pp 2854–2864
31. Wang M, Wang Y (2020) A synset relation-enhanced framework with a try-again mechanism for word sense disambiguation. In: Proceedings of the 2020 conference on Empirical Methods in Natural Language Processing (EMNLP), pp 6229-6240
32. Yin S, Zhou Y, Li C, Wang S, Ji J, Chen X, Wang R (2019) KDSL: a knowledge-driven supervised learning framework for word sense disambiguation. In: 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, pp 1–8
33. Patil A, Patil C, Ramteke R, Bhavsar RP, Darbari H. Different approaches of Marathi language word sense disambiguation. Int J Emerging Trends Eng Basic Sci (IJEEBS). ISSN (Online): 2349-6967
34. Patil A, Ramteke R, Bhavsar RP, Darbari H. Marathi language word sense disambiguation using modifies lesk algorithm
35. Patil A, Ramteke R, Bhavsar RP, Darbari H. Graph-based algorithm for word sense disambiguation: a performance and comparison
36. Zobaed S, Haque ME, Rabby MF, Salehi MA (2021) Senspick: sense picking for word sense disambiguation. In: 2021 IEEE 15th International Conference on Semantic Computing (ICSC). IEEE, pp 318–324

Impact of ICT in Education: An Analysis

Golak Bihari Palai and Deepanjali Mishra

Abstract One day in the month of November, 2019, the world received a major setback when it understood that a new pandemic called COVID-19, or the Novel Coronavirus had taken over to create havoc among the people. It was first started in a wet market of a small province in China. After that it has spread all over the world like a bonfire. Many countries got under its grip, namely USA, South Korea, and Italy where the situation was totally out of control. During the last two years, India suffered huge loss due to COVID-19 in terms of life, property, and other assets. India is the second largest most populous country with 130 crores was highly affected due to COVID-19. Out of all the aspects, education was the worst affected sector. There was no option left but to implement e-learning as a methodology of teaching and learning. It has emerged as one of the major sources of business, like e-commerce, learning methodology, and e-learning. E-learning is a methodology of teaching and learning where the teacher teaches using multimedia, and the learner learns using the digital mode of education. This mode of teaching and learning has indeed brought a revolution in the education process because neither the teacher nor the student needs to be together in one place. There are numerous subjects which can be taught online, ranging from technical to non-technical subjects. Literature is an imitation of fiction or non-fiction. Online could be the best mode of instruction for literature students. Therefore, this paper is an attempt to make an analysis of the implications of e-learning in education, and its implementation to teach literature by the teachers.

Keywords Literature · Digital media · E-learning · Multimedia · Fiction

1 Introduction

"Each improvement in correspondence from the print machine to the PC has delivered new social impacts," said Jefferey Cole, chief of the Center for the Digital Future, based at the USC Annenberg School for Communication and Journalism. "Presently,

G. B. Palai · D. Mishra (✉)
School of Humanities, KIIT University, Bhubaneswar, India
e-mail: Deepanjalimishra2008@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_45

as intuitive portable innovation progressively turns into an all day, every day experience, we discover clients remember they are receiving new kinds of rewards from innovation—alongside negative impacts at home and work." [1]. The words bode well in light of the fact that online innovation has in fact acquired an unrest different ramifications of innovation. Out of the entirety of its uses, encouraging procedure utilizing on the Web strategies has end up being a lot of compelling and prevails with regards to having different effect in the realm of innovation. The model connections educators to their students and to proficient substance, assets, and frameworks to assist them with improving their own guidance and customize learning. Web-based learning talks about the utilization of open instructive assets and different advancements that can be incremented instructively and profitability by quickening the pace of picking up; lessening costs related with instructional materials or program conveyance; and better using educator time [2].

Teaching literature has become a challenging task in recent times. Once upon a time, there was a flair for literature and people irrespective of age group used to experience the beauty of their lives only through poetry, novels and prose, and plays. For example, Charles Dickens' The Pickwick Papers were sold at a massive 35,000 per month on an average basis. It was not just the literate people who preferred to read Dickens, but there are instances where even illiterate people loved to visit the informal reading sessions in casual places like barber shops, hotels, etc., which clearly meant that Dickens was undisputed literary giant. Applebee has talked about conducting survey nationwide in order to determine the book length in schools, and he was quite successful in it [3]. J. Thompkins, whose reader response theory is widely popular, too can be accessed through online database [4].

NEA, an agency, conducted a survey on the literary habits of reading by people in the USA in the year 2002 and opined that: "Reading a book requires a degree of active attention and engagement. Indeed, reading itself is a progressive skill that depends on years of education and practice. By contrast, most electronic media such as television, recordings, and radio make fewer demands on their audiences, and often require no more than passive participation. Even interactive electronic media, such as video games and the Internet, foster shorter attention spans, and accelerated gratification. To lose such intellectual capabilities—and the many sorts of human continuity it allows—would constitute a vast cultural impoverishment." [5].

1.1 Artificial Intelligence (AI)

Artificial Intelligence or AI, as it is more often addressed, has played an important role in the societal development in employment, ethics, and education which has benefited various sections of the society. It is used in the following spheres in the field of education.

Communication: It enables instant communication between students and teachers so that they are able to get connected with other forms of AI globally. Differentiation:

Its due to AI, teachers and the students are able to access their resources that are available so that they can use it when needed [6].

Personalizing Learning Opportunities: Artificial intelligence is more helpful to more personalize learning opportunities for students and also to analyze the student responses, which determine the areas of need and interest, new resources could be determined which facilitates the students in order to understand the content in a much better way.

Assessments: Artificial intelligence enables the evaluator or a trainer assess the assignments and activities of the students in a much easier way without much efforts and thus making the work more easy. It also saves lot of time as it enables the teacher to assess the assignments together at the same time.

1.2 Big Data

Big data brings about a revolution in teaching technologies. It provides a base for interaction between trainers, learners, and administrators It provides an interface in the educational institutions between the students and the faculties by bringing the views of the students from their perspectives to the classes that they attend which are taken by the faculty members [7]. Apart from that, big data ushers in latest technological innovations bring in huge impact on the education system providing new opportunities to many users which is increasing everyday.

1.3 Avenues of Research

With the advancement of data collection and analytics methods, administrators will have more information to enhance probabilities of revolutionary enrollment in the future and to take decisions regarding enrolling the new students which facilitates educational institution drive for its growth and development. Through proper utilization of resources which is beneficial not only for students but also for facilitators [8].

1.4 Faculty Development

Big data facilitates development of digital literacy among the faculty members which is a promising strategy and seen as a very much progressive and beneficial. The faculties can create their own customized curriculum using the resources that are available online like e-learning materials, e-books, and school software programs [9]. Apart from that, automation enables the user can use the tools and create their assignments,

activities, etc., and manage their own learning management which enables the faculty to devote more time for their students and give them their maximum output.

1.5 Cloud Computing

Cloud computing is a platform where resources, applications, storage, and information can be shared over a specific network. There are two types of cloud computing, namely front-end and back-end. The front-end is the user interface, whereas the back-end section provides a platform of application and infrastructure of the cloud environment. It allows users to access applications, resources, and data stored in the cloud environment. Front-end interface is that application, at the same time, back-end interface is the major component of cloud computing technology that secures the information. It is the station of all the servers, computers, and databases [10]. There is a main central server which carries out few functions for the operations to take place. The functioning of cloud computing takes place through work shift of some important workload of resources. However, in case of a local computer, there is no need for the user to use many resources when it comes to running applications. Rather it is performed by. The network of servers in the cloud environment resulting in a substantial decrease in the software and hardware resources of the user.

1.6 Blockchain

Blockchain is a type of shared database among the nodes of a computer network in a distributed form. Maintaining records are one of the major functions of blockchain in education. There are various constrains of a faculty or a trainer who has to assess numerous records of the students that is time-consuming, endless, and also hectic. Lots of documentation is involved like assessing, evaluating, and grading. However, blockchain makes it easier, and it can perform the tasks of more than 738,000 unique credentials documenting learning can be counted, including degrees, certificates, digital badges, and apprenticeships. It helps in transfer of files online and by this blockchain, it is possible to check the records of the students by himself/herself and verify their courses which is assigned to them. Similarly, this facility can be availed by employees working in an organization.

1.7 5G

5G network has indeed brought about a revolution in the telecommunication and Internet of Things. It is 5th Generation technology in cellular networks. India has emerged as one of the major countries depending on cellular communication for

its daily needs. It has technologically advanced the people and has made them more techno savvy. Indians contribute and connect themselves to various activities related to technology. The present Government in India has taken initiatives and tried to implement certain schemes ensuring students to furnish them with an aim at enhancing their futurist skills. But unfortunately the ongoing pandemic COVID-19 disrupted its schemes and plans to a greater extent. Due to its student outbreak, people were confined to their homes and had to resort to online education. As the students were from versatile background related to social economical status or geography, many students were unable to access this mode of education due to poor network connectivity or unavailability of resources. But 5G network helped to solve major problems. It has been adopted by many schools and colleges across the country which facilitates in e-learning, advanced technology, and providing computer facilities to all. With each passing day, it is increasingly being adopted by countries and educational institutions all over the world. It provides smart classrooms which imparts better learning tools having positive outcome. This makes education system more effective and vulnerable. 5G enables the learners to be creative and innovative by communicating with their fellow students globally. It enhances the critical thinking skills of the students and make them collaborate with others in a much more impressive way.

2 An Approach to Teach Literary Criticism Through Virtual Mode

It is really challenging for a teacher of literature to teach using multimedia. In the words of Bolter and Landow, "numerous contemporary researchers accept that sight and sound are unmistakably appropriate for reaction-based ways to deal with the instructing, learning, and evaluation of abstract comprehension yet such ideas presently can't seem to be melodically investigated." The progressing "Sight and sound and Literature Teaching and Learning Project" at the National Research Center on Literature Teaching and Learning is more concerned about investigating the traits of multimedia which aids in the development of an abstract or while teaching comprehension. A significant target of the undertaking was to create rules to support instructors and designers consider interactive media from such a point of view. These models were utilized to survey business media writing applications and their job accordingly based educating and learning. It was more preferable to use multimedia during teaching as it would be easy to detect the response of the students. A trainer has to consider so many factors while using online methodology for taking a class of literature [11]. First of all, it is considered to be an outdated topic, and sometimes this subject is scheduled in the last period which makes the student bunk the class or even if the student is present, it is very difficult to make them concentrate on the topic.

One of the prime objective of a faculty of literature is to facilitate the learners to study or to look for the literary works online. He or she can instruct them to perform activities either on individual basis or in a group depending on the text. It is comparatively easy to locate information, get the question and answer, links, and content in a very effective manner if the students get their own computer which provides ample opportunities to the learners who can easily retrieve lot of relevant materials in the form of classroom notes, question, and answer in the form of pdf or ppt or word document. The facilitator could create assignment or send the classroom notes to the students which enables the learners to access in a very independent manner. Technology-enabled communication could intensify the literary texts by converting the literal hypertexts into any convenient form. One of the basic positive aspects of these hypertexts is that they are effective in nature and are responsible in stimulating ideas and motivate teamwork. Online training does not only contain the study materials or assignments which are given by the faculties, but it also includes student engaging activities to be performed in groups like discussions working on movie reviews and translation work by the learners themselves. In fact when the students opt for online classes to study literature, it is very necessary that they should use hypertext because it provides ample opportunities and scope for the students to explore its vast scope. Students can easily connect with the users of the outer world and other literary communities and research groups. Technological advancement has led to formation of major tools by the 2.0 Web version as it is easy to form social collaborative platforms and create videos [12]. Social media is fast emerging to be one of the motivating factor of the students, and it also makes teachers convenient to take their classes. Both of them could feel their presence in the magical world of literature. Ann Matthews Woodlief—a professor at the English Department at Commonwealth University of Virginia, USA, in one of her publications addresses the topic of teaching literature by using technology. “Pedagogically”, she explains, “my goal is to use a computer system to help create more dynamic communities of readers and literary translators, to develop such class lessons that allow each student to feel involved in intensive reading/thinking/writing activities and what is more to make it part of the larger readers’ and translators’ communities.” It has been found out that after the end of the course, the students feel that they have not only learnt more, but rather they have understood how to work on their assignments. They further opined that the classes taken online was much more interesting than the previous ones. It gave them a creative mind which would comprehend things in a more practical way.

3 Digitalization of Texts

Apart from this, Internet also enables in accessing various works of mythology and folklore online which benefits the students, researchers, or even many interested people. It even facilitates lots of people to take a virtual tour of their favorite

pilgrimage destinations so that they can have darshan of their favorite deity which saves time, money, and energy [13].

4 Advantages of Teaching Literature Through Multimedia

Teaching literature using online methodology has its advantages both for the teachers as well as the students. It enhances the studying and learning process of the students, and it can be beneficial to them in generating interest for literature. The multimedia strategy is rich wellspring of showing assets both in viable applications, and furthermore, it is considered beneficial to show assets which cannot just give essential language learning materials to students yet in addition offer multi-practical showing asset recovery for educators, to address different issues of instructors in the showing procedure, for example, looking for an article or planning questions, etc. The subsequent bit of leeway is the individualization of showing process: The corpus-based versatile learning mode accomplishes customized language learning through the development of corpus. It can cause the trouble of instructing materials to adjust to the language level of the student and the substance of training materials meet the understudy's adapting needs [14]. All the more significantly, the learning framework is perfect with simple language lab instructing, so it can additionally improve students' language learning results and premiums.

The self-versatile learning model of multimedia-aided teaching framework has a positive significance to progress of showing proficiency, change from the educator driven showing mode and students' free learning mode to the understudy controlled autonomous mode, and change from "coddling" example to students' self-enlistment. In this manner, it can in a roundabout way improve students' self-learning capacity and environment in a work pace of English alumni, getting a charge out of a high research worth and wide application prospect for the change of English instructing. In any case, the development of corpus identified with English learning isn't great. Furthermore, it is as yet important to continually investigate and find many research bearings of the interactive media showing self-versatile mode [15]. A few investigations have recommended that understudy fulfillment and inspiration are higher in courses that utilization sight and sound materials. A study analyzed the mentalities of more than 700 undergrads toward the utilization of PC innovation in twenty courses speaking to a wide scope of scholarly trains. Students were commonly positive about the utilization of innovation, in spite of the fact that females evaluated the utilization of innovation for learning, writing, and homeroom guidance fairly lower than did their male friends. Notwithstanding this, every person is armed up for the new innovation. It is imperative to remember that an inadequately created as well as executed utilization of mixed media can accomplish more mischief than anything. The potential instructive worth and reason for utilizing study hall media in these three focuses are as per the following:

To raise interest level:

- Students acknowledge (and regularly anticipate) an assortment of media.

To upgrade understanding:

- Rich media materials support understudy appreciation of complex points, particularly powerful procedures that unfurl after some time.

To build memorability:

- Rich media materials lead to better encoding and simpler recovery.

5 Conclusion

Therefore, it can be concluded that multimedia learning could be an effective alternative to teach humanities subjects like literature and language. In this manner, e-learning as a methodology for imparting education makes the door open for students to improve their learning and style for writing adequately [15]. This multimedia mode of learning will permit the students to decide their very own qualities and shortcomings and gain from them. Instructors can easily combine the multimedia into their homeroom by recognizing the learning styles of every one of their students, coordinating encouraging techniques to students' sight and sound learning for troublesome assignments, reinforcing more vulnerable students' interactive media learning through simpler undertakings and drill, and showing students, determination of learning methodologies. The day is not far behind when faculties of humanities would prefer to teach through online methodology over physical classroom.

References

1. https://news.usc.edu/54256/the-benefits-and-drawbacks-of-online-technology/. Retrieved 15 May 2020
2. https://www.ed.gov/oii-news/use-technology-teaching-and-learning. Retrieved 15 May 2020
3. Applebee AN (1989) A study of book-length works taught in high school English courses (report series 1.2). National Research Center on Literature Teaching and Learning, State University of New York, Albany
4. Thompkins J (ed) (1980) Reader-response criticism: from formalism to poststructuralism. Johns Hopkins University Press, Baltimore
5. Webb N (1983) Microcomputer learning in small groups: cognitive requirements and group processes. J Educ Psychol 76(6):1076–1088
6. Bolter JD (1991) Writing space: the computer, hypertext, and the history of writing. Erlbaum, Hillsdale. Retrieved 15 May 2020
7. Steve (1989) Another World: inside artificial reality. PC Comput 2(11):90–99, 102. Retrieved 15 May 2020
8. Bennett S, Lockyer L (1999) The impact of digital technologies on teaching and learning in K-12 education. Curriculum Corporation. Retrieved from: http://www.ndlrn.edu.au/verve/_resources/impact_digital_technologies_k-12pdf.pdf
9. Boruszko G (2013) New technologies and teaching comparative literature. CLCWeb Comp Lit Cul 15(6), Article Number 3. https://doi.org/10.7771/1481-4374.2355

10. Porter S (1999) Introduction: technology in teaching literature and culture: some reflections. In: Teaching European literature and culture with communication and information technologies. http://users.ox.ac.uk
11. Pratt MK (2009) How technology is changing what we read. Retrieved from http://www.pcworld.com/article/164355/e_books.html
12. Rozema RA (2004) Electronic literacy: teaching literary reading through the digital medium. Retrieved from http://scholarworks.wmich.edu/dissertations/1300/
13. Bauerlein M (2008) Online literacy is a lesser kind. The Chronicle Review, The Chronicle of Higher Education, 19 Sept 2008. http://chronicle.com/article/Online-Literacy-Is-a-Lesser/28307
14. Bauerlein M (2008) The dumbest generation: how the digital age stupefies young Americans and jeopardizes our future (or, don't trust anyone under 30). Jeremy P. Tarcher; Penguin, New York, p 110
15. (2010). http://lists.digitalhumanities.org/pipermail/humanist/2010-April/001183.html (my emphasis)

Identification of Inter-ictal Activity from EEG Signal Using Scalograms with LeNet-5 Based Model

Arshpreet Kaur and Kumar Shashvat

Abstract Identification of inter-ictal activity has always presented as a diagnostic challenge, for neurologist consuming much of their time. The automation of the process can provide the required support to the neurologist. Publically available Bonn data dataset has been used for this work. We have created two second segments of public data and created its scalogram which acts as an input to our model, whereas earlier researchers have worked on complete 23.6 s data. LeNet-5-based model is used as classifier**.** The goal of this work is to distinguish inter-ictal activity with and without presence of various artifacts. Accuracy of 98.03% has been accomplished for the public dataset.

Keywords Scalogram · LeNet-5 · EEG · Epilepsy · Classification

1 Introduction

Epileptic seizures are triggered by a disruption in the electric action in the brain. Epilepsy is a neurological disorder where patients may suffer spontaneous seizures. As proposed in [1], the significant characteristic of epilepsy is recurrent seizures. Figure 1 shows different types of seizures in epileptic seizures.

Many researchers have used different techniques to identify seizures using IEEG data such as in [3] and EEG data to detect seizures such as in [4, 5]. Different features have been explored by researchers to understand the dynamics of EEG, and various studies investigate different linear and nonlinear signal properties. Among which nonlinear properties have been deployed at large. Tzimourta et al. [6] and Thilagaraj et al. [7] used entropy-based features. Different discriminative models have been deployed by these researchers, support vector machine (SVM) and artificial

A. Kaur (✉) · K. Shashvat
Department of Computer Science & Engineering, Alliance College of Engineering and Design, Alliance University, Bengaluru, India
e-mail: arshpreet@nitdelhi.ac.in

K. Shashvat
e-mail: shashvat.sharma@alliance.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_46

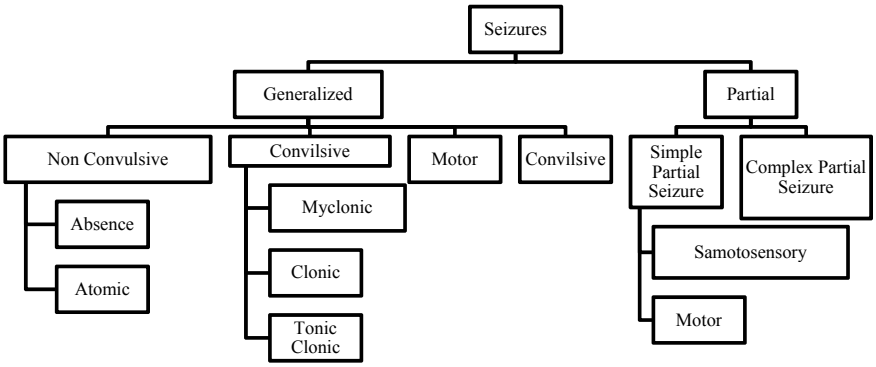


Fig. 1 Types of seizures [2]

neural network (ANN) being the most common ones. Even the recent researches conducted between 2018 and 2020 studies [8–11] have also deployed the SVM model or its variants for different case considerations; other models such as *K*-nearest neighbors (KNN) and PNN (probabilistic neural network) by researchers such as in [12]. Much effort has been devoted to develop epileptic EEG classification techniques. In Jerger et al. [1], the work analyzes consequences of seven straight and nonlinear techniques (examination of intensity spectra, cross-relationship, head parts, stage, wavelets, connection basic, and shared expectation) in recognizing the most punctual dynamical changes going before 12 intracranial-recorded seizures from 4 patients. Similar work can be found in [13–16].

2 Methodology

The entire procedure is separated into five components. The procedure is embodied in Fig. 2. Step 1, is considering data from the benchmark data, the details of which are mentioned in Sect. 2.1, under step 2 , the complete signals into segments of two seconds. Further in step 3 scalogram are formed using continuous wavelet transform. In step 4 the model is applied and under step 5 results are discussed.

Fig. 2 Work methodology

Table 1 Details of benchmark dataset

Total data folders	Five
Time	23.6 s
Sampling frequency	173.6 Hz

Table 2 Data details and group division

Healthy and inter-ictal	
Case 1	A-CD
Case 2	AB-C
Case 3	AB-CD
Case 4	AB-D
Case 5	B-CD

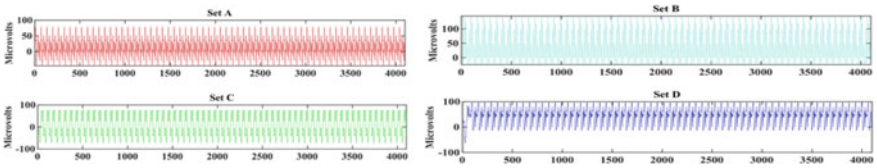


Fig. 3 Plots of different sets used

2.1 Data from Participants

2.1.1 Public Data

Figure 1 shows the pictorial representation of Bonn data, used in this work. The details are presented in Table 1. Table 2 holds the information of cases considered in this work. Figure 3 shows the plots of signals that have been used in this work.

In this work, our focus has been to identify inter-ictal discharges, and hence, we have used five cases from this group.

2.2 Data Division

The complete signals were split into smaller signals of two seconds.

2.3 Development of Scalograms

Figure 4 represents the scalogram created for four sets considered in this work.

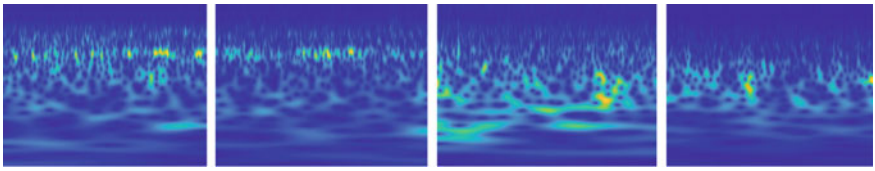


Fig. 4 Scalogram representation

2.4 *Resizing Images*

The dimension of the original scalogram achieved was 656 × 875 × 3, which was resized to 32 × 32 × 3.

2.5 *LeNet-5 Model*

LeNet-5 [1] was used in this work. Figure 4 describes its layout and Table 3 details the parameters used in it. Figure 5 depicts the LeNet-5 model diagrammatically.

Table 3 Parameters of the model

Layer	Filter size	Size of feature map	Activation function	Filters
Input		32 × 32 × 3		
Convolution layer	5 × 5	28 × 28 × 6	Tanh	6
Sub sampling layer	2 × 2	14 × 14 × 6		
Convolution layer 2	5 × 5	10 × 10 × 16	Tanh	16
Subsampling layer	2 × 2	5 × 5 × 16		
Convolution layer	5 × 5	120	Tanh	20
Fully connected		84	Tanh	
Fully connected		2	Softmax	

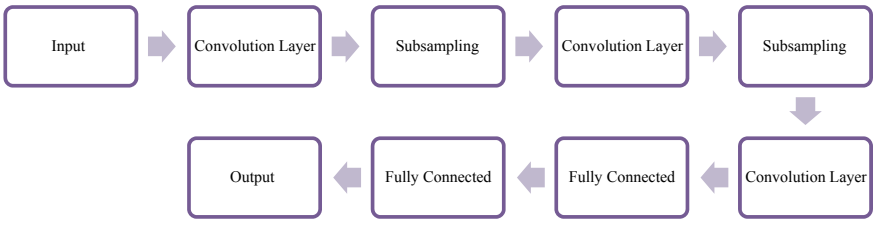


Fig. 5 LeNet-5 model description

Table 4 Hyperparameters used

Parameter	Value
Learning rate	0.01
Epochs	10
Shuffle	Every-epoch
Batch-size	32
Validation frequency	32

2.6 Hyperparameters

Table 3 describes the parameters used and Table 4 details the hyperparameters used in the network.

2.7 Validation Method

On the complete dataset, *k*-fold validation is employed, and the value of *k* is placed to ten.

2.8 Evaluation Parameters

The output from the model obtained is the class label. The predicted output is compared with the actual labels to form a confusion matrix. Accuracy, precision, and recall were obtained from the matrix.

3 Results

Intel® core™, i7-7700 CPU was used for this work. Random access memory of 16 GB was deployed and 64-bit operating was used. MATLAB 2018 was the programming platform used for this work. Table 4 holds the results of the cases considered. For the evaluation of the proposed model, we considered three performance evaluation parameters accuracy, specificity, and sensitivity. Table 5 presents the results obtained for (Case 1–Case 5).

Table 5 Results obtained from the proposed method

Cases	Accuracy (%)	Precision (%)	Sensitivity (%)
A-CD	96.21	92.95	96.20
AB-C	95.60	97.21	92.00
AB-CD	95.79	96.19	95.71
AB-D	97.12	97.89	96.00
B-CD	98.03	96.02	98.25

Table 6 Comparison of performance evaluation parameters for group 5

Author	Year	Cases	Sensitivity (%)	Specificity (%)	Accuracy (%)
Siuly et al. [17]	2010	A-D	89.9	80	84.9
Jaiswal and Banka [18]	2016	A–D	99.85	99.95	99.5
Hekim [19]	2016	AB–CD	93	91	92
Sriraam [20]	2017	A-C	99.4	100	99.7
Manish et al. [21]	2017	AB-CD	90.5	94.5	92.5
Proposed work	2022	AB-CD	95.79	96.19	95.71

4 Discussion

In this study, five cases have been considered, out of which reporting of only one cases is available in previous researches, i.e., case 3 (AB-CD). For the other cases, which are case 1, case 2, case 4, and case 5 that have not been previously reported, we report classification accuracy of 96.21%, 95.79%, 97.12%, and 98.03%, respectively. In context to sensitivity and specificity in Table 6, our method has performed better than all the methods compared. Table 6 presents the studies that have been conducted on similar lines, and it is observed that all the studies used complete length signals of 23.6 s and did not present results for all the cases considered under this work. However, the presented model still performs better than the compared work for most of the cases.

5 Conclusion

It is evident from this work that using the current methodology where scalogram is used as input, with existing models such as LeNet. For the purpose of inter-ictal identification for all the five cases, the algorithms have performed proficiently. In this work, only public data has been used, and the use of real-time data could enhance the knowledge about the model use in such settings. Also, for the future work, comparisons with more diverse convolution network models could be made.

Acknowledgements The authors convey admiration to the public accessible database [22].

References

1. Jerger KK, Weinstein SL, Sauer T, Schiff SJ (2005) Multivariate linear discrimination of seizures. Clin Neurophysiol 116(3):545–551
2. Cavanna A, Monaco F (2009) Brain mechanisms of altered conscious states during epileptic seizures. Nat Rev Neurol 5:267–276
3. Varatharajah Y, Iyer RK, Berry BM, Worrell GA, Brinkmann BH (2017) Seizure forecasting and the preictal state in canine epilepsy. Int J Neural Syst 27(1):1–21
4. O'Shea A, Boaz AL, Chambers M (2019) A hierarchy of power: the place of patient and public involvement in healthcare service development. Front Sociol 4(May):1–12
5. Fu R, Han M, Tian Y, Shi P (2020) Improvement motor imagery EEG classification based on sparse common spatial pattern and regularized discriminant analysis. J Neurosci Methods 343(February):108833
6. Tzimourta KD, Tzallas AT, Giannakeas N et al (2019) A robust methodology for classification of epileptic seizures in EEG signals. Health Technol (Berl) 9:135–142
7. Thilagaraj M, Rajasekaran MP, Arun Kumar N (2019) Tsallis entropy: as a new single feature with the least computation time for classification of epileptic seizures. Cluster Comput 22:15213–1522
8. Sharmila A, Suman AR, Pandey S, Mahalakshmi P (2018) Epileptic seizure detection using DWT-based approximate entropy, Shannon entropy and support vector machine: a case study. J Med Eng Technol 0:1–8
9. Ramakrishnan S, Muthanantha Murugavel AS (2019) Epileptic seizure detection using fuzzy-rules-based sub-band specific features and layered multi-class SVM. Pattern Anal Appl 22:1161–1176
10. Li M, Chen W, Zhang T (2019) FuzzyEn-based features in FrFT-WPT domain for epileptic seizure detection. Neural Comput Appl 31:9335–9348
11. Yu J, Wang L, Chen X (2019) Epileptic seizure classification based on the combined features 0–5
12. Goshvarpour A, Goshvarpour A (2020) Diagnosis of epileptic EEG using a lagged Poincare plot in combination with the autocorrelation. Signal Image Video Process
13. O'Regan S, Faul S, Marnane W (2013) Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals. Med Eng Phys 35(7):867–874
14. Berg P, Scherg M (1991) Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG. Clin Phys Physiol Meas 12(A):49–54
15. Croft RJ, Barry RJ (2000) Removal of ocular artifact from the EEG: a review. Clin Neurophysiol 30(1):5–19
16. Gasser T, Ziegler P, Gattaz WF (1992) The deleterious effect of ocular artefacts on the quantitative EEG, and a remedy. Eur Arch Psychiatry Clin Neurosci 241(6):352–356; 8(2):24–28
17. Siuly S, Li Y, Wen PP (2011) Clustering technique-based least square support vector machine for EEG signal classification. Comput Methods Programs Biomed 104:358–372. https://doi.org/10.1016/j.cmpb.2010.11.014
18. Jaiswal AK, Banka H (2017) Biomedical signal processing and control local pattern transformation based feature extraction techniques for classification of epileptic EEG signals. Biomed Signal Process Control 34:81–92. https://doi.org/10.1016/j.bspc.2017.01.005
19. Hekim M (2016) The classification of EEG signals using discretization-based entropy and the adaptive neuro-fuzzy inference system. https://doi.org/10.3906/elk-1306-164

20. Sriraam SRN (2016) Classification of epileptic seizures using wavelet packet log energy and norm entropies with recurrent Elman neural network classifier. Cogn Neurodyn. https://doi.org/10.1007/s11571-016-9408-y
21. Sharma M, Pachori RB, Rajendra Acharya U (2017) A new approach to characterize epileptic seizures using analytic time-frequency flexible wavelet transform and fractal dimension. Pattern Recognit Lett 94:172–179. https://doi.org/10.1016/j.patrec.2017.03.023
22. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE (2001) Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev E Stat Nonlin Soft Matter Phys 64(6 Pt 1):061907

Stock Market Price Prediction Using Machine Learning and Deep Learning Techniques

Shivani Raina, Adwait Kadam, Pratik Sawant, and Minal Apsangi

Abstract Stock expense assessing is a notable and a huge point in money-related and academic assessments. Protection trade is an unsteady spot for expecting since there are no tremendous standards to survey or anticipate the expense of the stock in the protections trade. Various strategies like particular examination, fundamental assessment, time series examination and quantifiable assessment, etc., are used to predict the expense in tie monetary trade; yet, these strategies are by and large not exhibited as a dependably OK assumption instrument. In this paper, we done a long-short term memory (LSTM) method for managing and expect stock market costs. LSTMs are really completed in deciding stock expenses, returns, and stocks showing. We outline the arrangement of the LSTM with its noteworthy features and versatile limits. We revolve around a certain social occasion of limits with a fairly basic impact on the stock expense of an association. By taking assistance of assessment, it was discovered that the limit marks of the new article and that supported gauging definite result. Though monetary trade can never be expected with hundred percent precision due to its equivocal space, this paper targets exhibiting the capability of LSTM at assessing the stock expenses.

Keywords AI · Deep learning · TensorFlow · Stock market · ML · LSTM

1 Introduction

The financial market is a mind-boggling framework where individuals can trade stocks, values, monetary forms and different wares over virtual stages upheld by specialists.

The financial exchange permits the financial backers to possess portions of public organizations. The market has given financial backers and brokers monstrous measures of profits over few introductory ventures and furthermore gave okay

S. Raina (✉) · A. Kadam · P. Sawant · M. Apsangi
PVG's COET GKPIM, Pune, India
e-mail: shivaniraina777@gmail.com
URL: https://www.pvgcoet.ac.in/

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_47

contrasted with the gamble of beginning another business or going for more lucrative profession choices.

Throughout the long term, financial backers and brokers have been chipping away at creating and testing the frameworks to investigate the way of behaving of the securities exchange. Notwithstanding, examining the market developments and conduct is incredibly difficult on the grounds that the securities exchange is exceptionally unstable and profoundly interrelated with factors like financial, political mental and area explicit factors. Technical and fundamental analysis are the two main approaches taken to analyse the stock market. To invest or trade and to increase the chances of high profits with low risk, these two major approaches are used to make decisions in the market by traders and investors.

The primary objective of the paper is to study and apply profound learning strategies to the securities exchange to anticipate the value conduct of the stock and subsequently follow up on those forecasts to decrease the gamble of venture misfortune and create benefit. The forecasts will be tried against the noteworthy information. This venture will be useful to fledgling brokers to go with better choices.

To do as such, this paper will utilize the best-fitting model, for example one which yields the best outcomes; for this situation of time series anticipating, it will be the LSTM model that represents long short-term memory. Contrasted with an ordinary profound brain organization, its viability is because of the expansion of a critical part in time series forecast, the memory part.

2 Literature Review

In this part, we investigate ongoing exchanging models in light of AI. The development of endlessly stock cost file for Indian financial exchanges was shown in our most memorable paper [1]. He differentiates four expectation models in his exploration, with two techniques for input to these models. The main methodology incorporates costs for OHLC, while the subsequent methodology centers around the deterministic pattern results. Results show that three other expectation calculations are performed by random forest classifier for the main methodology. The random forest methodology was presented by Kompella and Chakravarthy examined other boundary classifications, with genuinely enormous consequences for an organization's portion cost [2]. Study showed the extremity score of new items anticipated exact results. Tae, Joon, Deuk, So Young presented measurements of the monetary organization that could be applied to speculation procedures for worldwide capital business sectors [3]. They proposed worldwide financial exchange directed and roundabout instability networks in view of basic pairwise relationship and framework-wide availability of public stock lists utilizing an auto-backward model of the vector. In light of their exhibition examination information, network pointers have been demonstrated to be significant extra markers in worldwide expectation. During market emergencies, the market and those pointers were more fruitful. Our next examination is centered around [4]. This exploration utilizes diagram hypothesis to propose a

clever methodology. This method fetches information concerning the spatio-transient connection between different stocks by displaying the securities exchange as an intricate organization.

3 Methodology

3.1 Data Collection

Data comprises of profits, for example rate change in close cost of various organizations isolated into little, normal and huge capitals money as in Chart 1. An organization's market capitalization is the stock valuation. Little cap is a term used to recognize the exceptional protections in nearly little market capitalization organizations. Ordinarily an organization underneath Rs. 5000 crores in India is considered as a little cap business. In little cap organizations, one can foresee genuinely high unpredictability.

3.2 Data Exploration

After reading the data from Yahoo Finance, we collected it into a data frame (Fig. 1).

We want data with patterns occurring over time. The graph below already says a lots of things. This chart is overflowing with various ways of behaving of financial exchange over time. This will make the learning more powerful as well as gave us a change to test how great the forecasts are for an assortment of circumstances (Fig. 2).

	Date	Open	High	Low	Close
0	1970-01-02	0.30627	0.30627	0.30627	0.30627
1	1970-01-05	0.30627	0.31768	0.30627	0.31385
2	1970-01-06	0.31385	0.31385	0.30996	0.30996
3	1970-01-07	0.31385	0.31385	0.31385	0.31385
4	1970-01-08	0.31385	0.31768	0.31385	0.31385

Fig. 1 Format of the dataset

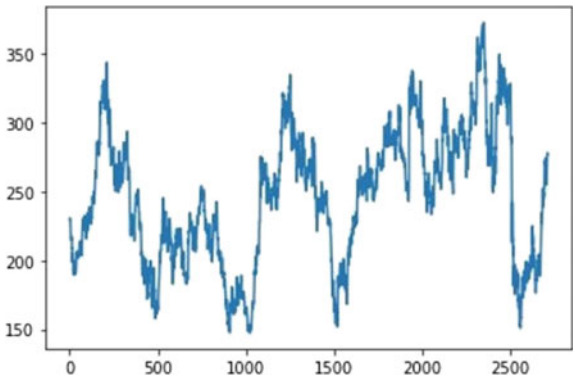


Fig. 2 Graph for closing price versus days

3.3 Data Normalization

Presently, we characterized a scaler to standardize the information. MinMaxScalar scales every one of the information to be in the locale of 0 and 1. We likewise reshaped the preparation and test information to be in the shape. Because of the perception we made before, that is to say, unique time spans of information have different worth reaches, we standardize the information by parting the full series into windows. On the off chance that we did not do this, the prior information will be near 0 and would not enhance the growing experience.

While picking the window size, we ensured it is not excessively little, since when we perform windowed-standardization, it can present a break at the finish of every window, as every window is standardized freely.

3.4 Data Augmentation

To make our model vigorous, we did not make the result for x_t consistently x_{t+1}. We actually think this is a sensible presumption for stock development expectations.

Beneath we showed how a group of information is made outwardly (Fig. 3).

3.5 Defining Hyperparameters

We characterized a few hyperboundaries. D is the dimensionality of the information. It is direct, as we accept the past stock cost as the info and anticipate the following one, which ought to be 1. Then, we have num_unrollings; this is a hyperparameter connected with the backpropagation through time (BPTT) that is utilized to improve the LSTM model. This means the number of ceaseless time steps we consider for a solitary streamlining step. We can consider this; rather than improving the model

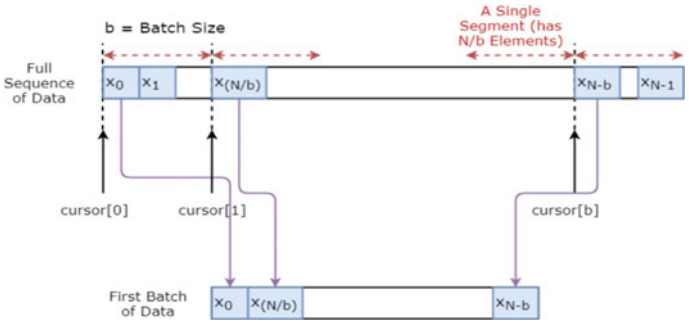


Fig. 3 Illustration of a batch of data

by taking a gander at a solitary time step, we upgrade the organization by seeing num_unrollings time steps. The bigger the better. Then, we have the batch_size. Bunch size is the number of information tests you consider in a solitary time step.

4 Proposed System

Long short-term memory models are very strong time-series models. They can anticipate an inconsistent number of steps into what has to come. A LSTM module (or cell) has five fundamental parts which permits it to display both long haul and momentary information.

Cell state (ct)—This addresses the inside memory of the cell which stores both transient memory and long-haul recollections.

Hidden state (ht)—This is yield state data determined with respect to current information, past secret state, and current cell input which we in the end use to anticipate the future securities exchange costs. Also, the secret state can choose to just retrieve the short or long haul or the two sorts of memory put away in the cell state to make the following forecast.

Input gate (it)—Decides how much data from current information streams to the cell state.

Forget gate (ft)—Decides how much data from the ongoing information and the past cell state streams into the ongoing cell state.

Output gate (ot)—Decides how much data from the ongoing cell state streams into the secret state, so that if necessary LSTM can pick the drawn out recollections or transient recollections and long haul recollections.

Also, the conditions for working out every one of these substances are as per the following (Fig. 4).

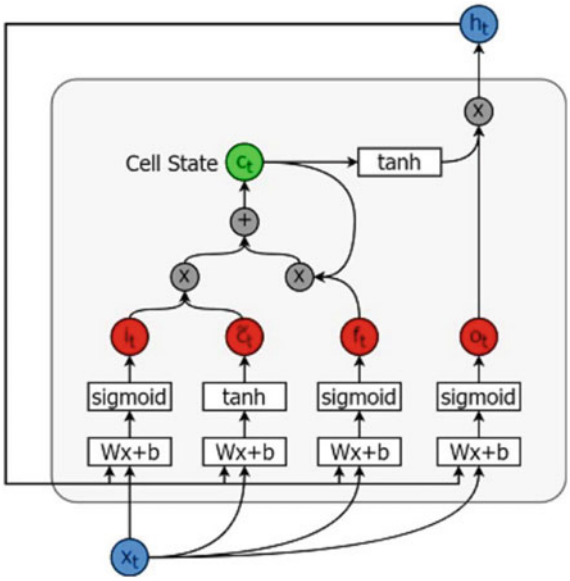


Fig. 4 Cell in LSTM model

it = σ(WixXt + Wihht-1 + bi)
c~_t=σ(W_cxx_t+W_chht−1+bc)
f_t=σ(W_fxx_t+W_fhht−1+bf)

ct=f_tc_t−1+itc~t
o_t=σ(W_oxx_t+W_ohht−1+bo)
ht=ottanh(ct)

5 Experiments and Results

Subsequent to running the LSTM model, the MSE misfortune was going down with how much preparation. This was great sign that the model was getting the hang of something helpful. To evaluate our discoveries, we can contrast the organization's MSE misfortune with the MSE misfortune we got while doing the standard averaging. We can see that the LSTM is improving (Figs. 5 and 6).

However somewhat flawed, LSTMs appear to be ready to foresee stock cost conduct accurately more often than not.

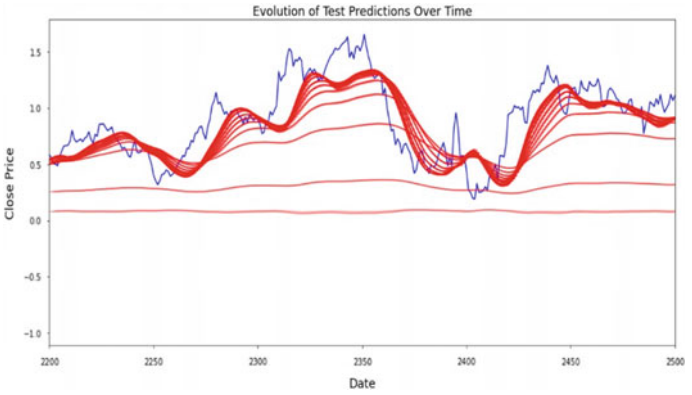


Fig. 5 Evolution of test prediction over time

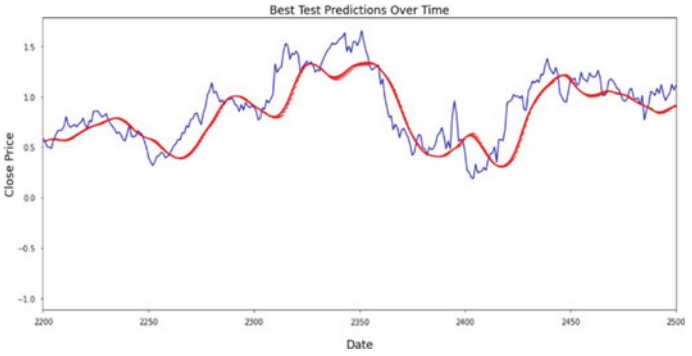


Fig. 6 Best test prediction over time

6 Conclusion

The created application does its forecast of the end stock cost of some random association, with a base measure of blunder. The utilization of LSTM calculations in stock forecast has shown to be fundamental and has in this manner denoted their utilization in solid market plans. Such methodologies would be an extraordinary assistance for dealers and financial backers to put cash in the securities exchange since they depend on a massive scope of authentic information in the wake of being minded example information. Enhancement of additional applicable sources of info, for example principal information and specialized information from subordinate business sectors, might work on the consistency of the model. Later on, we can expand this application for foreseeing digital money exchanging and furthermore, we can add feeling examination for better forecasts; likewise, future works for this task will incorporate and require more inside and out research with an end goal to adjust the best outcome yielding model, the half and half model to the Indian market.

Henceforth, assessing the results, it can be implied that innovation additionally has ample distance to go before it is totally fit for settling the securities exchanges' secret.

References

1. J P (2020) Stock market analytics: statistical and machine learning techniques. Int J Psychosoc Rehabil 24(1):1828–1833
2. Kompella S, Chilukuri K (2019) Stock market prediction using machine learning methods. Int J Comput Eng Technol 10(3)
3. Lee T, Cho J, Kwon D, Sohn S (2019) Global stock market investment strategies based on financial network indicators using machine learning techniques. Exp Syst Appl 117:228–242
4. (2019) Prediction of option price using ensemble of machine learning algorithms for indian stock market. Int J Rec Technol Eng 8(2):3231–3241
5. Shah D, Isah H, Zulkernine F (2019) Stock market analysis: a review and taxonomy of prediction techniques

Design and Analysis of the Radiation Characteristics of a Microstrip Antenna Array Resonating at 28 GHz for Satellite Applications

Padire Mounika, T. Jaya, and B. Ebenezer Abishek

Abstract A microstrip antenna array with a resonant frequency of 28 GHz for satellite applications is built, and its radiation characteristics are investigated in this research. The design is a microstrip patch antenna array fed by a corporate network. The Feko Suite was used to examine the radiation properties of the samples. The design was optimised with the help of the optimisation feature in Feko Suite in order to get better radiation properties. The need for satellite applications to operate in the 28 GHz frequency spectrum is also emphasised. An in-depth discussion of the design of a microstrip antenna array as well as the explanation of the results is provided.

Keywords Gain · Microstrip array · Return loss · Series feed · Resonant array · 28 GHz

1 Introduction

In recent years, there has been a significant increase in the number of satellite applications. Telecommunications is the application that accounts for the majority of satellite usage. Other key uses include weather forecasting, navigation, remote sensing and other similar tasks. Consumer demand for bandwidth has increased as a result of the increase in the number of consumers. Bottlenecks and congestion issues have arisen as a result of the expanding use of satellite technology at lower frequency bands. As a result, satellite applications now necessitate the use of a greater frequency spectrum than previously required.

P. Mounika (✉) · T. Jaya
VELS Institute of Science, Technology and Advanced Studies (VISTAS), Chennai, India
e-mail: padire.mounika@gmail.com

T. Jaya
e-mail: jaya.se@velsuniv.ac.in

B. Ebenezer Abishek
Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Chennai, India
e-mail: ebenezerabishek@veltechmultitech.org

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_48

The use of Internet of Things technology, as well as the rise of smart city initiatives, will result in a multiple-fold increase in traffic. It is a well-known truth that spectrum is a limited resource; therefore, a proper technique for reusing spectrum should be employed in order to maximise the efficiency with which spectrum is utilised.

Massive MIMO is a new technology that combines hundreds of antennas to transmit multiple messages at the same time in a small amount of space. Despite the fact that hundreds of antennas are employed, they are all phase synced, allowing the data signals to be combined together before reaching the target user.

The satellite's footprint is extensive, and it is the most appropriate option for covering the vast geographical area that encompasses India. The cost of land is extremely high in most parts of India, especially in the urban areas. As opposed to terrestrial networks, satellite broadband will not necessitate the acquisition of extensive territory. The same spectrum can be reused by the satellite in geostationary orbits, resulting in high-speed services at a reasonable cost throughout the nation. Satellites with ultra-high-throughput (UHTH) capabilities can help reach the goal of providing countrywide access. Because of the large footprint of the satellite, it is possible to contact billions of people simultaneously throughout the entire country. These ultra-high-throughput satellites are capable of providing both fixed fibre broadband services and mobile connections similar to 5G for consumers on the move. The millimetre-wave spectrum (28 GHz) is required for the improved connection provided by satellite broadband. Satellite broadband is the most effective solution for dealing with the problem of limited fibre availability. The entire cost of deploying the broadband network will be decreased as well. As a result, rural India will be able to benefit from the network without the requirement for cable laying, and the entire deployment process will be expedited because the slow pace of cable laying will not be required. All citizens, regardless of their geographical location, can take advantage of the cost-effective satellite broadband in all of their gadgets.

Antennas for satellite applications, which are widely employed, command a lot of attention. The performance of microstrip antennas is excellent, but the cost and space required are considerable; as a result, we can utilise them as a low-profile alternative to high-profile antennas despite their superior performance. There are a large number of satellites orbiting the planet, each of which provides a variety of commercial services to people all over the world. Ku and C bands are the ones that are most commonly used for satellite applications, but the Ka band is emerging as the one that is best suited for satellite applications, and it has the potential to provide a solution to consumer demand as well as congestion difficulties in the lower frequency bands. The design and analysis of a low-profile resonant series-fed rectangular microstrip patch antenna array are presented in this study, as well as their radiation characteristics [1].

The length of feed lines has been lowered, the efficiency of the array has been increased, and the complexity has been reduced when compared to corporate-fed microstrip patch arrays. Microstrip patch arrays with series feed are hence the ideal choice for satellite applications.

Arrays that resonant with each other and arrays that travel with the waves are the two types of series-fed patch arrays. In travelling wave arrays, the size of the wave

transmitted from the series feed diminishes over time, and only the residual power is received by the load as a result. The patch width is responsible for controlling the intended distribution of excitation. The progressive phase change between the radiating parts allows for the generation of a broadside fan beam and the scanning of a beam. It is also possible to produce better matching of the antenna over a wider bandwidth, in contrast to resonant series-fed patch arrays that have open circuit termination, resulting in the activation of standing waves at the patch edges. The bandwidth can also be increased by using better design [2, 3].

2 A Review of the Literature

The future 5G communication systems are expected to significantly increase communication capacity by utilising vast unlicensed bandwidth, notably in the millimetre-wave range, which is currently underutilised. Also projected to be ready for and capable of supporting extremely high data rates, which will present a major challenge in terms of network requirements as well as antenna designs in order to meet the anticipated data rate and capacity [4–6]. Because of the evolution of wireless communication systems, low-profile antenna types that are capable of giving astounding performance throughout a wide frequency band are required to be developed. Due to their low fabrication cost, lightweight and small volume and low-profile layout when compared to other bulky types of antennas, the MSPA is an excellent choice for wireless devices in this regard. The MSPA is simple to use and versatile in terms of polarisation, resonant frequency, pattern and input impedance, among other characteristics. They can be fitted to the surfaces of high-performance aeroplanes, space crafts, rockets, satellites, missiles, automobiles and even hand-held mobile cell phones to provide improved reception. As a result, the MSPA plays a significant role in the wireless communications industry, which is now experiencing the strongest growth. By enhancing surface wave and spurious feed radiation laterally across the feeding line, the depth of the substrate material, on the other hand, degrades the MSPA bandwidth and radiation efficiency, as shown in Fig. 1. As a result, feed radiation effects guide undesirable cross-polarised radiation in a certain direction. Losses such as conductor, dielectric, and radiation are experienced by the MSPA as a result of which the bandwidth is narrowed and the gain is decreased [7–13]. As a result of this performance limitation, the bandwidth of MSPA is limited; its directivity, gain, and radiation efficiency are all low, making it unsuitable for usage in the next-generation 5G communication systems. There have been a number of various designs demonstrated and described by scientists in the scientific literature in an attempt to improve MSPA performance for 5G communication. Many broadband patch antenna designs are presented in [7–9, 14–16] to boost the bandwidth and radiation efficiency of the antennas used in these applications. This includes the patch with a multi-layer substrate integrated waveguide, and multi-patch designs that incorporate several slots on the patch, as well as employing a defective ground plane, tweaking the width of the patch, and using a serial feed to connect multiple patches together.

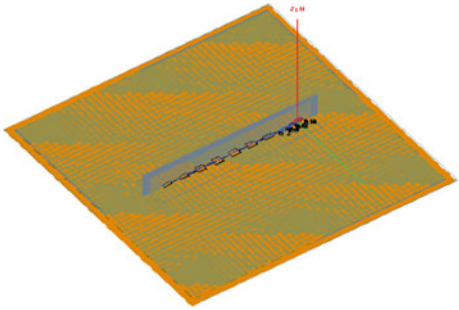


Fig. 1 Series-fed resonant microstrip antenna array

In a similar vein, the usage of quarter-wavelength microstrip feed line as inset-feed and lumped element to the patch edge were proposed in [10, 15, 16] to reduce the feeding network topologies and patch edge impedance mismatch. According to the above-mentioned methodologies, the rectangular patch antenna provided in [19–21, 24] achieves minimum return losses and a wide −10 dB impedance bandwidth when operated at the resonant frequency. Alternatively, the antenna return loss described in [6, 8, 18, 22] is significant for wireless communication because of poor impedance matching at the interfaces. Furthermore, while the antenna published in [18–23] has a high beam gain, the design reported in [6, 10, 22, 24] achieves a relatively narrow bandwidth, which is insufficient for 5G applications. Another MSPA of a similar design is provided in [6, 8, 22, 23], which has a lower VSWR and higher radiation efficiency than the previous one. As a result, based on the above summary of previously acquired simulation findings in earlier work, we can infer that the exhibited work is an attempt to improve functionality in terms of one or two specific performance metrics through simulation. Specifically, we are motivated by the desire to develop and analyse a microstrip antenna operating at 28 GHz in order to determine its practicality for satellite systems while also increasing all of the antenna's essential performance parameters. In order to do this, we have used inset-feed and quarter-wavelength impedance matching techniques, as well as modifying the antenna's size.

3 Materials and the Methodology

It is the shape of the antennas' physical geometry, the dimensions of their antenna structures, and the material properties from which they are constructed that determine their performance characteristics the most significantly. The rectangular patch shape is used in this study because it is simple to design and analyse, and it has a wide bandwidth because it is broader in shape when compared to other types, making it an excellent choice. The physical structure of the MSPA under investigation is depicted in Fig. 2. The MSPA is made of the FR-4 substrate material, which has a loss tangent of 0.0025, a r of 4.4, and a radiating copper metal thickness of 0.035 mm. It is

intended to operate at a frequency of 28 GHz and has a loss tangent of 0.0025. The bandwidth, directivity, return loss, gain and radiation efficiency of the MSPA are all characteristics that describe its performance. In order to achieve our goal, we have employed techniques such as inset-feed impedance matching, quarter-wavelength impedance transformer, and antenna dimension optimisation as part of our methodology. A table listing the various physical dimensions of the planned antenna that were determined for the radiation of antenna with a resonance frequency of 28 GHz is provided in Table 1.

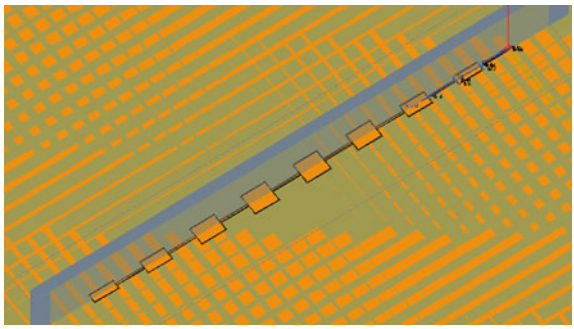


Fig. 2 Radiating patch elements of series-fed resonant mirostrip antenna array

Table 1 Physical dimensions of series-fed resonant mirostrip antenna array

Parameter	Dimension
Centre frequency	28 GHz
Number of elements	8
Length of the feed line	3.994 mm
Width of the feed line	343.6 mm
Length of the patch	3.570 mm
Width of the patch 1	1.328 mm
Width of the patch 2	1.894 mm
Width of the patch 3	2.618 mm
Width of the patch 4	3.077 mm
Width of the patch 5	3.077 mm
Width of the patch 6	2.618 mm
Width of the patch 7	1.894 mm
Width of the patch 8	1.328 mm
Substrate height	10 mils
Relative permittivity	2.2
Loss tangent	0

4 Result and Discussion

In this section, the simulation results and discussions of the proposed rectangular MSPA are presented, as well as its design and implementation. In order to analyse the designed antenna, we simulated the proposed MSPA design using the CST software package. Different performance metrics are employed in order to gain access to the characteristics of an antenna. The bandwidth, gain, directivity, VSWR, and return loss are among the parameters that are frequently used. The magnitude of the return loss between the radiating patch and the feed point is used to determine the quality of the matching between the two. The antenna is designed to work with a 50 Ω feed line. Figure 3 shows the three-dimensional gain plot of the series-fed resonant microstrip antenna array. The maximum gain of 20 dB is achieved at the resonating frequency of 28 GHz.

Figure 4 shows the E-plane pattern of the series-fed resonant microstrip antenna array. The maximum gain of 20 dB is achieved at the resonating frequency of 28 GHz.

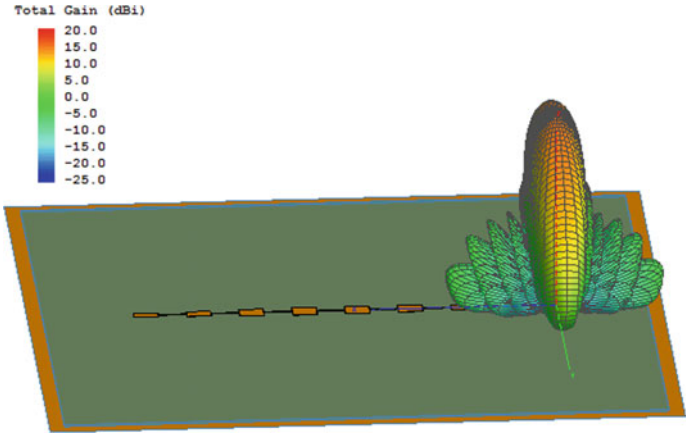


Fig. 3 3D gain plot

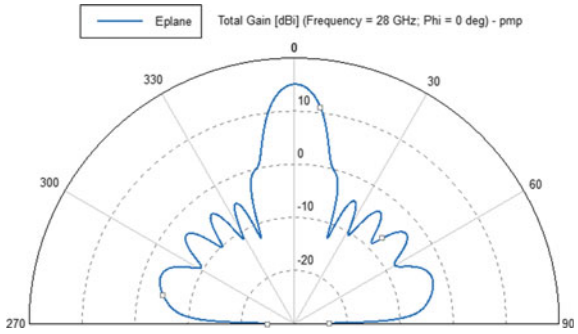


Fig. 4 2D gain plot E-plane pattern

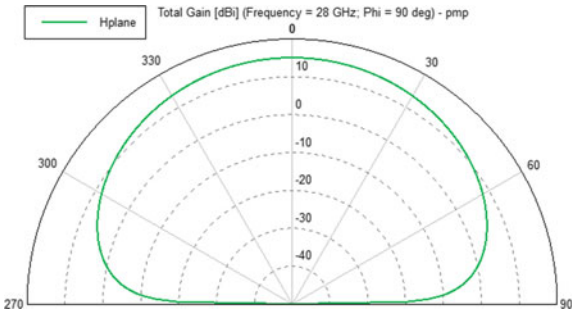


Fig. 5 2D gain plot H-plane pattern

Figure 5 shows the H-plane pattern of the series-fed resonant microstrip antenna array. The maximum gain of 20 dB is achieved at the resonating frequency of 28 GHz.

Figure 6 shows the return loss plot of the series-fed resonant microstrip antenna array. The maximum gain of 20 dB is achieved at the resonating frequency of 28 GHz. The impedance bandwidth of the designed antenna measured at −20 dB is 271 MHz.

The performance of the designed antenna is compared to that of existing similar designs in the literature, as shown in Table 2. According to the results of [10, 18–24], the proposed design outperforms the designs mentioned in those publications. Similar to this, the developed antenna exhibits superior performance in terms of radiation efficiency compared to the designs provided in [8, 10, 19, 20, 24]. When compared to the designs published in [8, 18, 22, 23], the VSWR of the examined MSPA is the smallest possible, which is significantly closer to optimum values. At 28 GHz, the investigated antenna exhibits the lowest return losses when compared to designs described in [8, 18, 19, 22–24], while it exhibits the highest return losses when compared to designs reported in [20, 21]. Lastly, when compared to the designs published in [10, 22, 24], the proposed MSPA achieves a wide bandwidth, but it is narrow when compared to the designs reported in [19–21]. As a result, when

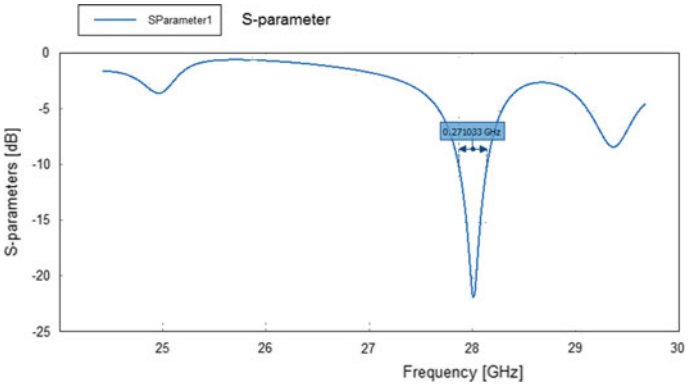


Fig. 6 Return loss of series-fed resonant mirostrip antenna array

compared to other similar antennas previously reported in the literature, the suggested patch antenna exhibits extremely competitive characteristics.

Table 2 Charateristics of existing similar designs as per the literature survey

References	S_{11} (dB)	Gain (dBi)	VSWR	η_{rad} (%)	BW (GHz)
Yuan et al. [8]	−15.35	–	1.79	87.8	–
Braaten et al. [10]	−20.53	6.21	1.02	65.6	0.4
Goyal and Modani [18]	−17.4	6.72	1.28	–	–
Abdelsalam et al. [19]	−23.67	6.7	–	81.2	1.150
Kaeib et al. [20]	−39.37	6.37	1.022	86.73	2.48
Ghazaoui et al. [21]	−39.7	5.23	–	–	4.1
Kavitha et al. [22]	−14.151	6.06	1.488	–	0.8
Sivabalan et al. [23]	−22.2	6.85	1.34	–	–
Darsono and Wijaya [24]	−27.7	6.72	1.22	75.875	0.463
This work	−22.9	20	1.04	–	0.271033

5 Conclusion

The design and performance study of a rectangular MSPA operating at 28 GHz for satellite applications are discussed in this research. The proposed simulation result shows that the return loss, directivity, beam gain, and bandwidth are all − 38.86553 dB, 7.509 dBi, 7.587 dBi, and 1.046 GHz, respectively, with the return loss, directivity, beam gain, and bandwidth being −38.86553 dB, 7.509 dBi, 1.046 GHz, and the bandwidth being 1.046 GHz. When compared to existing designs that have been published in the scientific literature, the proposed antenna exhibits much superior performance. Improvements in performance have been accomplished in this article as a result of the introduction of a combined optimisation of the parameters, inset-feed impedance matching, and quarter-wavelength impedance matching. As a result, the antenna type proposed in this study is an excellent contender for millimetre-wave wireless applications in the satellite era of communication.

References

1. Alam MM, Sonchoy MMR, Goni MO (2009) Design and performance analysis of microstrip array antenna. In: Progress in electromagnetics research symposium proceedings, Moscow, Russia, 18–21 August 2009
2. Waterhouse R (2003) Microstrip patch antennas: a designer's guide, 1st ed. Springer
3. Haupt RL (2010) Antenna arrays: a computational approach. Wiley, Hoboken, New Jersey

4. Stutzman WL, Thiele GA (1998) Antenna theory and design, 2nd ed. Wiley, New York, NY. In: 2013 7th European conference on antennas and propagation (EuCAP) 2025
5. O'Donovan PL, Rudge AW (1973) Adaptive control of a flexible linear array. Electron Lett 9(6):121–122
6. Balanis CA (2005) Antenna theory: analysis and design, 3rd edn. Wiley, New York, NY
7. Caloz C, Itoh T (2005) Electromagnetic metamaterials. Wiley-IEEE Press, Piscataway Hoboken, NJ
8. Yuan T, Yuan N, Li L (2008) A novel series-fed taper antenna array design. IEEE Antennas Propag Lett 7:362–365
9. Lai A, Itoh T, Caloz C (2004) Composite right/left-handed transmission line metamaterials. IEEE Microwave Mag 34–50
10. Braaten BD, Roy S, Ullah I, Nariyal S, Ijaz B, Masud MM, Naqvi SA, Iftikhar A (2012) A cascaded reconfigurable RH/CRLHZero-phase microstrip transmisison line unit cell. In: Proceedings of the IEEE international conference on wireless information technology and systems (ICWITS), Maui, Hawaii, 11th–16th November 2012
11. Monti G, De Paolis R, Tarricone L (2009) Design of a 3-state reconfigurable CRLH transmission line based on MEMS switches. Prog Electro Res 95:283–297
12. Jang T, Hwang S-H, Bang Y-S, Kim J-M, Kim Y-K, Baek C-W, Lim S (2009) Switchable composite right/left-handed (S-CRLH) transmission line using MEMS switches. IEEE Microw Wirel Compon Lett 19(12):804–806
13. Hashemi M, Yuan N, Itoh T (2008) Dual-band composite right/left-handed metamaterial concept. IEEE Microwave Wirel Compon Lett 22(5):248–250
14. (2009) Advanced design system-ADS, vol 7. Agilent Technologies, CA
15. Dheeraj M, Shankar D (2018) Design and analysis of 28 GHz millimeter-wave antenna array for 5G communication systems. JSTEMARI 1(3):1–9
16. Kothari N, Sharma S (2018) A 28 GHz U-slot microstrip patch antenna for 5G applications. Int J Electromagn Dev Res (IJEDR) 6(1):363–368
17. Bergsrud C, Freidig C, Anderson M, Clausing M, Dito T, Noghanian S Inset-fed and edge-fed patch antennas with rectifying circuit
18. Goyal RK, Modani US (2018) A compact microstrip patch antenna at 28GHz for 5G wireless applications. In: IEEE conference record, third international conference, and workshops on recent advances and innovations in engineering, November 2018
19. Abdelsalam M, Omar M, Elgnai AM, Abuo EH (2019) Design and analysis of millimeter wave microstrip patch antenna for 5G applications. In: International conference on technical sciences, pp 137–142
20. Kaeib AF, Shebani NM, Zarek AR (2019) Design and analysis of a slotted microstrip antenna for 5G communication networks at 28 GHz. In: 19th international conference on sciences and techniques of automatic control and computer engineering (STA), March 2019
21. Ghazaoui Y, El Alami A, El Ghzaoui M, Das S, Barad D, Mohapatra S (2020) Millimeter-wave antenna with enhanced bandwidth for 5G wireless application. J Instrumentation (JINST)
22. Kavitha M, Dinesh Kumar T, Gayathri A, Koushick V (2020) 28GHz printed antenna for 5G communication with improved gain using array. Int J Sci Technol Res 9(3):5127–5133
23. Sivabalan A, Pavithra S, Selvarani R, Vinitha KM (2020) Design of microstrip patch antenna for 5G. Int J Control Autom 13:546–552
24. Darsono M, Wijaya AR (2020) Design and simulation of a rectangular patch microstrip antenna for the frequency of 28 GHz in 5G technology. In: International conference on innovation in research. J Phys Conf Ser 1469(012107):1–7. https://doi.org/10.1088/1742-6596/1469/1/012107
25. Ebenezer Abishek B, Raaza A, Ramesh S, Jerritta S, Rajendran V (2019) Circularly polarized circular slit planar antenna for vehicular satellite applications. Appl Comput Electromagn Soc J 34(9):1340–1345. SCI Indexed
26. Ebenezer Abishek B, Raaza A, Leelaram K, Jerritta S, Rajendran V Design and placement of a conformal microstrip antenna on a vehicular structure for SATCOM applications. In: Springer-lecture notes in networks and systems. https://doi.org/10.1007/978-981-10-8204-7_45. Scopus Indexed

27. Ebenezer Abishek B, Raaza A, Rajendran V (2018) Simulation analysis of circular and linearly polarized patch antenna for vehicular SATCOM. J Adv Res Dyn Control Syst 12(Special Issue):800–803. Scopus Indexed
28. Ebenezer Abishek B, Abdul Hashim Z, Prasath HS, Stephen CA, Kumar PCK Design of conformal microstrip patch antenna for vehicle tracking. Int J Eng Technol (UAE) 7(3.27):189–191. https://doi.org/10.14419/ijet.v7i3.27.17758, Scopus Indexed

A Survey: Extraction of River Networks from Satellite Images

Sukrut Bidwai, Devang Jagdale, Tejas Hiremath, Neil Bhutada, and Sukhada Bhingarkar

Abstract River network extraction is crucial to keep track of the water resources. Various methods have been implemented in times series to yield profound and incisive outputs and are still being developed and combined with predefined available methods. We have carried out a structured survey on these methods and have presented them with their outputs. There are numerous Web sites available for data set collection. Some generalized methods are available like image processing, using predefined models, or developing user-defined algorithms. For image processing, various segmentation methods are available out of which clustering- and threshold-based segmentations are mostly used. Predefined models such as CNN, ResUNet, YOLO, Faster CNN, and MSCFF are available. These algorithms can be used for the extraction of river networks but might not yield higher accuracies. Hence, this paper concentrates mainly on approaches for the extraction of river networks from satellite images.

Keywords CNN · Deep learning · DEM · D8 algorithm · Image processing · River network extraction · Mask generation · Satellite images · Segmentation

1 Introduction

Water resource is the most significant source of life on Earth. Water resources monitoring, protection, and flood management are predominant factor. Extraction of water bodies from a satellite image is a crucial task. Previously, the man-made space probes were assigned to collect huge volume of data in the form of graphic pictorial representations of certain surfaces, which resulted in load archives of images which were acquired. However, now this gigantic amount of data is reshaped into valuable executable data. Estimation and analysis of surface changes are crucial

S. Bidwai (✉) · D. Jagdale · T. Hiremath · N. Bhutada · S. Bhingarkar
Dr. Vishwanath Karad MIT—World Peace University, Pune, India
e-mail: s.bidwai2000@gmail.com

S. Bhingarkar
e-mail: sukhada.bhingarkar@mitwpu.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_49

for a better understanding and management of processes that cause them. Various machine learning models such as ResNet, RCNN, YOLO, Faster CNN, and ResUNet are further used for extraction of river networks. Initially, the basic steps include pre-processing of images and converting them into uniform size for further training and feature extraction. Image spectrum and image enhancements are added in order to extract RGB and grayscale values which will be further trained to extract the exact river network from the provided image based on training and validation testing. Thus, from above research, some attributes and models are taken into consideration for incisive outputs.

2 Data Source and Study Area

The data sources which are available consist of USGS, NASA, Google Earth Engine [1], Kaggle, Copernicus, Bhuvan NSRC GOV, etc. There are various satellites which are available on the following Web sites, e.g.,: Sentinel-I [2], Sentinel- II [2], TERRA, LISS-III, etc. Images can be selected based on various features like using map tiles, interactive drawing, latitudinal, and longitudinal coordinates, etc. Some Web sites also provide random images from different parts of Earth. These regions include Heihe River Basin, KG River Basin [3] which emanates southern area of tablelands. There are some Web sites available which provide filters for satellite images like shadows and clouds.

2.1 Image Processing Techniques for Extraction of River Networks

Researchers proposed a method for river network extraction using image processing.

Process includes taking input as a Landsat satellite image and produces a river mask as output. Initially, after taking satellite image as input, enhancement using decorrelation stretching is performed. In this process, image colors are enhanced by stretching their RGB bands which causes pixels to lose their correlation with each other which in turn causes them to normalize and removes high correlation to produce images with higher color contrast. By doing this feature, discrimination process becomes much easier [1] (Fig. 1).

After performing enhancement, the next step is image segmentation. Image is first converted to a grayscale image where grayscale level ranges up to 256 levels. Image segmentation is the process where we split pixels into various groups. Different groups contribute to different objects which make similar pixels fall under one object and exhibit a common meaning. Image segmentation makes the image simpler which helps us to analyze the image at a much lower complexity [1].

There are four different types of segmentations which are available as follows:

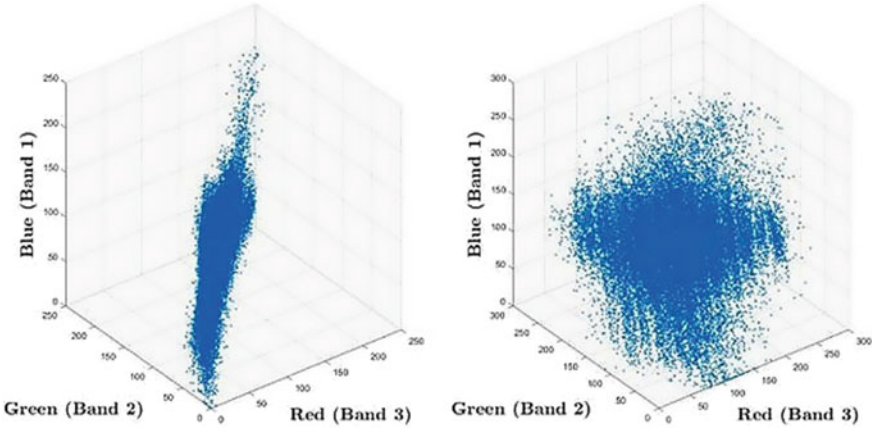


Fig. 1 Color scatterplot before **a** and after **b** decorrelation [1]

(a) Threshold segmentation
(b) Edge segmentation
(c) Region segmentation
(d) Clustering segmentation.

Threshold segmentations are compared with the intensities of each pixel with a given threshold and assign that pixel a group. There are two types of thresholds available:

(a) Global threshold: There is only one common threshold for whole image.
(b) Local threshold: There are multiple thresholds for different parts of images.

Global threshold method was better suited for rivers mask as we had to only group pixels into two groups. In threshold-based segmentation, threshold value has to be selected appropriately. Best possible value can be obtained by trial- and-error method, but instead of using an iterative method, we opt for histogram thresholding. Histogram thresholding considers that a given image has two parts: subject and background. It tries to find an optimal threshold value (T) so that the image can be divided into subject and background. If (image[x][y] <= T) assign group G_1 Else assign group G_2. In the above example, x and y denote intensity of pixel at given coordinates. In this way, we can assign pixel groups where G_1 and G_2 act as object and background. In clustering-based segmentation, we try to group pixels that are together. Pixel which is closer to each other is considered to be in relation and is grouped together. We can use K-means clustering for segmentation. In clustering, we try to find the cluster center so that pixels close to that center will be assigned a cluster. We keep repeating this process until we have only two clusters, subject and background [1].

Region-based segmentation is a repetitive process where pixels are grouped into smaller regions in each iteration to form a larger region. Initially, we define some seed points, and from these points, our region starts growing by claiming neighboring

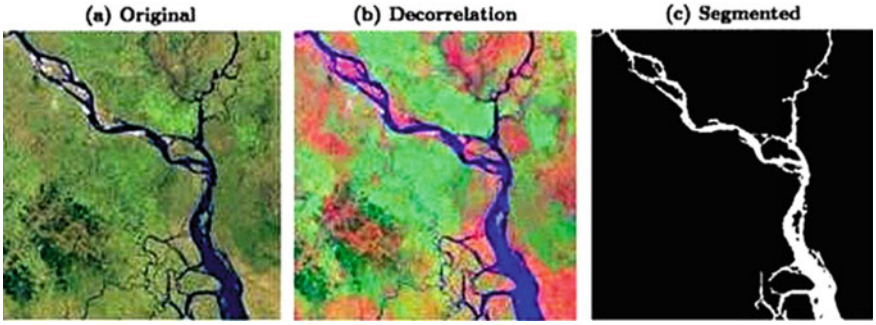


Fig. 2 Image segmentation for satellite images [1]

points which have similar properties like color, intensity, grayscale level, etc. This method has high time complexity compared to other segmentation methods, and it is also sensitive to noise.

Segmentation gives output as a binary image where the object is river (white) and background is land (black). Final step is to remove clutter from the image. After segmentation, the image might have noise or unwanted objects categorized as subject. To remove these unwanted objects from the image, we use morphological operations. In morphological operations, we compare a given pixel with its neighbor and decide whether to keep it or not. Process of removing pixels from an image is called erosion. In this process, we initialize the number of neighboring pixels a pixel should have. Pixel having fewer neighbors than selected value is transferred to another group, i.e., transfers pixel from subject to background. In this way, we obtained a river mask for a Landsat satellite image (Fig. 2).

The only limitation is that this method is not generalized. For images with different features like snow, clouds or shadows might result in false mask prediction.

2.2 CNN Method for Extraction of River Networks

The river networks can be extracted using few methods. One of them is using a predefined model, and second would be developing a new model based on any predefined model and using their outputs as new input to another model. Using a predefined model does not yield high accuracy at all times, so instead of using predefined models, we can combine different models to get high accuracy. One of the methods proposed includes a combination of multi-spectral data and synthetic aperture RADAR (SAR) data to jointly map the extent of the river over large regions. A neural network structure is defined to extract common information from multi-source data inputs, i.e., multi-spectral data and SAR (synthetic aperture RADAR) data [2].

A CNN is used to extract feature 1 and feature 2 from a network which is blocked and would be used during the model training and parameter transformations. In [2],

algorithms which have been defined include Algorithm 1 and Algorithm 2, whereas Algorithm 1 shows the pseudo-code for the model training. Here, the training set with two different types of the sample pairs are marked with a flag variable. Here, the flag variables are manually obtained during the multi-source sample pairs creation for common information extraction model training. The training sample pair with flag = 1 represents the relevant pair (a_i, b_i), and the training sample pair with flag = 0 represents the irrelevant pair (a_i, b_j). Given two input samples, Input 1 (e.g., a_i) and Input 2 (e.g., b_i or b_j), the loss function for extracting common information has been defined as

$$\text{Loss(Input1, Input2, flag)} = \text{Loss flag} = 0 + \text{Loss flag} = 1 \quad (2)$$

When flag = 0, the loss flag = 0 is computed as follows:

$$\text{Loss flag} = 0 = 0.5 * (1 - \text{flag})\{\max(0, m - \text{Distance})\}^2$$

In [2], Algorithm 1 defines the structure as the input acquired a sample pair of SAR and multi-spectral data, and a flag which is further defined as a multi-source training sample pair of relevant or irrelevant type. The output derived from this Algorithm 1 is the distance variable. Now, the feature 1 and feature 2 which are extracted from the CNN are blocked and the distance which calculated above is enabled with the combination of input 1 and input 2, i.e.: SAR and multispectral image sample input. The next step defined resembles of fine-tuning the network parameters using loss function Loss (Input 1, Input 2, flag) defined in equation. For each multi-source training sample pair, a new parameter is defined as m; now here, M is calculated as the Euclidean distance between input 1 and input 2 as given by Distance = $\|f(\text{input } 1)\|^2$.

In [2], Algorithm 2 defines the structure for testing the process of common information extraction function. Here, the feature 1 and feature 2 which were extracted using a CNN are enabled and the distance parameter is blocked. Input 1 and input 2 which are SAR data and multi-spectral data source, respectively, are considered as input. A common information matrix for multi-source testing sample pairs are produced from feature 1 and feature 2 for each multi-source testing sample pair (Figs. 3 and 4).

2.3 Mapping Network Using Geo-Processing

Digital elevation model consists of 3D characterization of elevation to represent terrain. It represents the fluctuating values between 2 points of elevation. The common use of DEM is to extract the terrain parameters for geomorphology. Three types of DEM (SRTM1-V3, AW3D, GDEM2, and ALOS PRISM) are used to compare the changes in morphology. The Alos prism module data is used to create the DEM. Demoted values are first identified and then filled in the DEM [3].

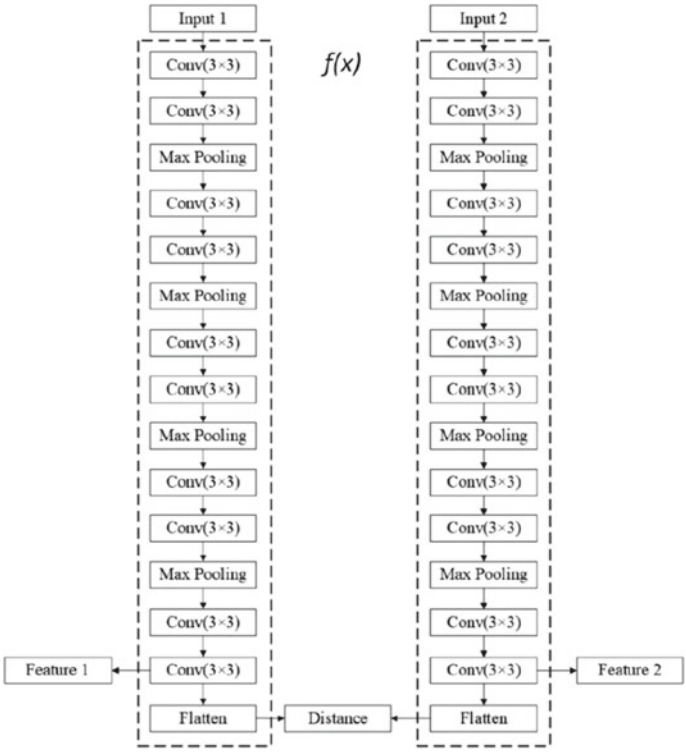


Fig. 3 CNN model for extraction of feature 1, feature 2, and distance parameter [2]

In [3], the D8 algorithm follows a basic principle that each cell or each grid can have eight different possible flows based on the maximum gradient. The D8 consists of mono-flow parameter algorithm and leads with a principle which states that only eight possible flows can be implemented in a single grid.

The algorithm exhibits further formula:

$$\text{Single Grid} = H \div C$$

where H = The height difference between the two cells and C = The distance between centers of two cells (Fig. 5).

The flow direction is determined using the principle of D8 algorithm. These directions are expressed by eight different signatures 128, 1, 2, 4, 8, 16, and 3. The D8 (eight direction) flow direction method assumes that water from any cell will flow to one of the steepest neighboring cells. After calculating the direction of flow from D8, furthermore calculation of the upstream confluence matrix is carried out, and then, model extracts the critical support areas from the raster map of classified water body network. Hence, raster diagram segmentation of the river network is extracted [3] (Fig. 6 and Table 1).

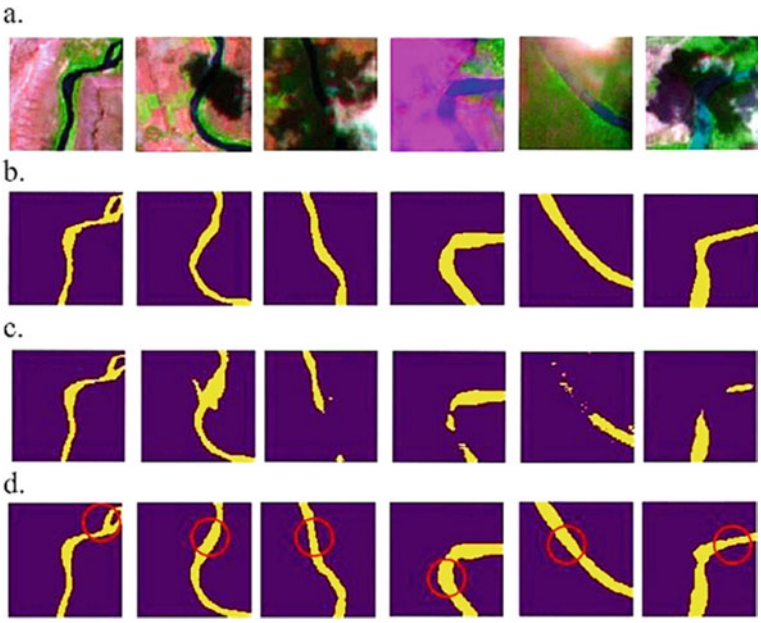


Fig. 4 Images in **a** are satellite images, images in **b** are ground truth, **c** output generated from predefined UNET model, and images in **d** are proposed method output [2]

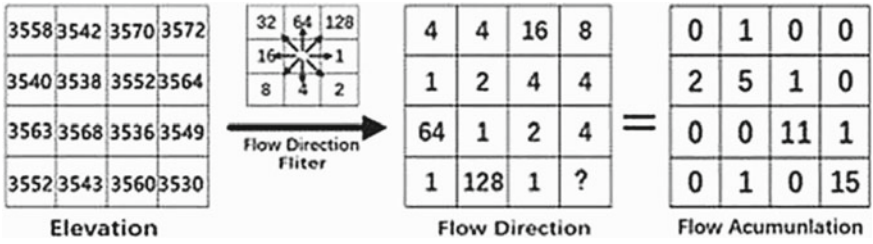


Fig. 5 Basic conceptual model of D8 algorithm [3]

3 Conclusion

Image processing gives accurately predicted river masks for satellite images. But this method is applicable for only one type of image. If the images contain a wide spectrum of filters like clouds, shadows, snows, etc., it does not adapt and applies an iterative method for all. If we have all images of the same kind, then image processing might be an appropriate choice. If we use predefined CNN models, we might get masks with lower accuracy as it is generalized and not specifically developed for river network extraction. But as future scope, we can develop ensemble predefined CNN models or tweak some of its parameters to get higher accuracy. Also, an additional benefit

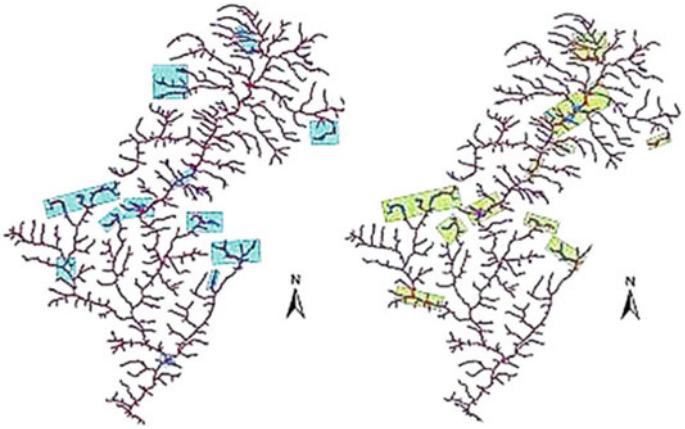


Fig. 6 Contrastive image of drainage system in SRTM1v3, ALOS3D, and ALOS prism algorithm [3]

Table 1 Summary of existing approaches

S. No	Methodology	Advantages	Limitations
1	Enhancement of images using the correlation, image segmentation, clutter removal [1]	It generates very high accuracy mask compared to predefined models	The parameter set which was used within the algorithm were not generalized for image segmentation and noise data elimination
2	Contrastive learning process to extract representative hidden features from multi-spectral data and SAR data [2]	Combined two different predefined models to produce river mask	2 datasets used, 1. Extremely low resolution images. 2. noisy dataset
3	The basic structure is divided into bi-parts: (a) Offline model training with MS CNN (b) Online prediction on Google Earth Engine [5]	Provides an executable deep learning models with respect to Google Earth Engine to improve the accuracy of river extraction	Image dataset was only of 36 images
4	Feature extraction using Gabor filter and canny edge detector are implemented for training and testing the random forest models [6]	Consistency is achieved which results in accuracy for network extraction	Only Canny's algorithm was used as an edge detection algorithm for feature extraction

(continued)

Table 1 (continued)

S. No	Methodology	Advantages	Limitations
5	A network is proposed, labeled as dense local feature compression (DLFC) network which aims to extract water body from different remote sensing images [7]	The experimental results are observed in which the model is perfectly able to distinguish between building and shadows (noise)	Distinctively small water bodies over certain set image dimensions are mixed in datasets
6	(CNN) is proposed, based on the multidimensional densely connected CNN for identifying water bodies in the lake area [8]	Optimal performance is achieved in identifying the boundaries and edges between land and water, whereas the mountain shadows and noise are distinguished	Optimal threshold to extract water varies with region and time
7	ResUNet is used as backbone in the architecture, along with Different post-processing methods like watershed methods [9]	Automatic river extraction is been done using predefined CNN algorithms	Dataset was collected in a small duration
8	Identification and removal of noise using suitable filters using k-means and region growing segmentation techniques for river extraction [10]	MSE are found to be lowest for k-means clustering algorithm compared to region- based clustering	Climatic conditions and shadows were not taken into consideration

of diverse image type can yield and support machine learning algorithms to yield a higher accuracy; in this way, we can use these methods for extraction of river networks to keep track of water resources through satellite imagery.

References

1. Ciaburri C, Benitez M, Sheta A, Braik M (2020) Automatic extraction of rivers from satellite images using image processing techniques. ACCENTS Trans Image Process Comput Vis 6:32–41. https://doi.org/10.19101/TIPCV.2020.618040
2. Wei Z, Jia K, Liu P, Jia X, Xie Y, Jiang Z (2021) Large-scale river mapping using contrastive learning and multi-source satellite imagery. Rem Sens 13(15):2893. https://doi.org/10.3390/rs13152893
3. Chen F, Tian B, Adhikari B, Gou X (2019) Mapping digital drainage network using geoprocessing: a case study of Kali Gandaki River Basin, Nepal Himalaya.https://doi.org/10.1109/IGARSS.2019.8900474
4. Yang X, Pavelsky T, Allen G, Donchyts G (2019) RivWidthCloud: an automated google earth engine algorithm for river width extraction from remotely sensed imagery. IEEE Geosci Rem Sens Lett 1–5. https://doi.org/10.1109/LGRS.2019.2920225

5. Asokan A, Anitha J (2019) Machine learning based image processing techniques for satellite image analysis—a survey. In: 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). https://doi.org/10.1109/comitcon.2019.8862452
6. Wang Y, Li Z, Zeng C, Xia G-S, Shen H (2020) An urban water extraction method combining deep learning and google earth engine. IEEE J Sel Topics Appl Earth Observ Rem Sens 13:768–781. https://doi.org/10.1109/JSTARS.2020.2971783
7. Naaz Z, Malini Devi G (2021) River network classification from multi-spatial satellite imagery using random forest. Int J Eng Res Technol (IJERT) 10(09)
8. Li M, Penghai W, Wang B, Honglyun P, Hui Y, Yanlan W (2021) A deep learning method of water body extraction from high resolution remote sensing images with multi-sensors. IEEE J Sel Topics Appl Earth Observ Rem Sens 1–1. https://doi.org/10.1109/JSTARS.2021.3060769
9. Wang G, Wu M, Wei X, Song H (2020) Water Identification from high-resolution remote sensing images based on multidimensional densely connected convolutional neural networks. Rem Sens 12(5):795. https://doi.org/10.3390/rs12050795
10. Waldner F, Diakogiannis FI (2020) Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network. Rem Sens Environ 245:111741. https://doi.org/10.1016/j.rse.2020.111741
11. Deepika RGM, Kapinaiah V (2017) Extraction of river from satellite images. In: 2017 2nd IEEE international conference on recent trends in electronics, information and communication technology (RTEICT). https://doi.org/10.1109/rteict.2017.8256591

UpAIsthiti: A Touchless Attendance System

Dimple Nachnani, Salonee Velonde, Sejal Kriplani, Mayur Pawar, Shashikant Dugad, Gresha Bhatia, and Abha Tewari

Abstract Managing attendance is a vital task for every institution. Considering the COVID pandemic where many organizations have resorted to online mode of working, it has become imperative to maintain social distancing and digitize various processes. Thus, for maintaining attendance of the students of schools/colleges or employees of a company, a touchless attendance system is required that records the attendance by capturing faces and does not waste time. This one-of-a-kind application uses a client–server model and captures the faces of students/employees through video feeds from mobile phone cameras, and the images are sent to a server, where image processing is used to process the faces. Further, with the help of dlib and the face recognition library, it identifies the faces and records the attendance in the software itself. The processed image is again sent back to the client android application, and the user gets notified about their attendance. Additional functionalities for data analysis and updating data have also been added to the system. Thus, the whole attendance system is an effort to make the attendance activity easy and efficient.

D. Nachnani (✉) · S. Velonde · S. Kriplani · M. Pawar · G. Bhatia · A. Tewari
Vivekanand Education Society's Institute of Technology, Mumbai, India
e-mail: 2018.dimple.nachnani@ves.ac.in

S. Velonde
e-mail: 2018.salonee.velonde@ves.ac.in

S. Kriplani
e-mail: 2018.sejal.kriplani@ves.ac.in

M. Pawar
e-mail: 2018.mayur.pawar@ves.ac.in

G. Bhatia
e-mail: gresha.bhatia@ves.ac.in

A. Tewari
e-mail: abha.tewari@ves.ac.in

S. Dugad
Tata Institute of Fundamental Research, Mumbai, India
e-mail: shashi@tifr.res.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_50

Keywords Touchless attendance system · Face detection · Face recognition · Authentication · Verification · Client–server model · Encryption · Socket programming

1 Introduction

Computers are becoming powerful enough to perform repetitive tasks previously performed by humans. With the growth of computing power, programs that use a variety of complex algorithms to perform activities that humans can no longer undertake 24 h a day, seven days a week can now be developed. Due to the ongoing pandemic, employees' function in the workplace has evolved, and safety measures such as social distancing and sanitization are used to provide a safer, healthier, and disease-free environment. Touchless attendance systems are a key aspect of resuming work in the middle of a pandemic.

Manual and fingerprint attendance techniques are a time-consuming process [7]. Users should be able to mark their presence without physically coming into contact with any device or person, which necessitates the usage of touchless attendance monitoring systems. Touchless attendance systems typically involve face attendance systems, radio frequency identification (RFID) attendance systems, GPS attendance systems, and iris identification attendance systems.

A unique radio frequency tag is assigned to each user in the RFID attendance system [10], and anytime the user presses their ID card against the scanner, it searches the database for that particular tag and records the attendance [8]. Facial encodings are stored in the database, and attendance is entered after scanning the user's face in the face attendance system.

Users can easily trick RFID [9] and GPS attendance systems, and as a result, they are ineffective. Since iris scanners are more expensive due to their high accuracy and reliability, they are not feasible for small institutions As a result, face recognition systems are cheap, appropriate, and an area of future development.

2 Related Work

The field of face recognition based on machine learning is growing rapidly, and a lot of work is being done to improve the accuracy of the systems implementing it. This section includes a review of the related methodologies employed for the same.

The groundwork for face detection has been laid in the paper [1] by the authors Viola and Jones who proposed the use of an integral image for quick calculations of features and an AdaBoost classifier to select the important ones and then combined the classifiers in a cascade to discard the background regions and focus on the facial features. This Haar cascade classifier detected faces 15 times faster than previously

existing systems. However, it faced various issues such as variation in illumination, the angle of the face made with the camera, and the scaling factor.

In paper [2], the authors have reviewed different face detection techniques based on feature-based and image-based approaches. The feature-based approach extracts the features of the face and builds a model to relate the features to detect a face, whereas in the image-based approach, the general characteristics of a face are learned from images of faces and non-faces. The feature-based approach is relatively easier to implement but is susceptible to differences in illumination, noise, and occlusion.

The authors Bah et al. in paper [3] have proposed to increase the quality of the characteristics extracted from an image, advanced image processing techniques such as contrast correction, histogram equalization, and so on are used. The faces are then recognized using a robust local binary pattern (LBP) face identification; however, the system is plagued by occlusion.

In paper [4], a deep learning neural network-based technique for face identification is proposed, which uses a cascade of six CNNs for face detection and FaceNet for mapping the features of the face to a 128-vector Euclidean space. A linear support vector machine (SVM) is used for classification purposes. The accuracy of the system is affected by the lighting conditions.

Principal component analysis (PCA) is a popular technique for detecting patterns in images in order to implement face recognition in paper [5]. Face detection is implemented using the Haar cascade classifier mentioned in [1], followed by PCA. The PCA algorithm is used to calculate eigenvalues for every face and store them in an XML file. The eigenvalue of the test face is calculated, and a match is found to the closest neighbor. The face detection rate decreases dramatically as the angle of the face with the camera increases.

3 Proposed Solution

In this paper, a client–server-based attendance system, UpAIsthiti, has been proposed for automating the process of recording attendance and making it completely touchless. The system makes use of a platform-independent Python-based server which is responsible for processing the images and marking the attendance. The client is an Android device that captures the live camera feed and sends it to the server for face recognition. At the administrator's end, the attendance can be recorded using two mobile devices for entry and exit, respectively. The faces are detected in the live camera feed using the face detection algorithm provided in the following section. These located faces are then encoded using the face-encoding algorithm and compared with the existing faces stored in the database. If no match is found, the person is deemed to be a visitor. For every face detected in the entry feed, the person's credentials along with the entry time are recorded in the database. Similarly, when a user is detected on the exit feed, the attendance is updated by adding the exit time.

The algorithms used in the system are as follows:

3.1 Face Detection

Face detection is the process of determining whether or not a human face is present in a given image. This is done by invoking the face locations function of the face recognition library, which accepts as parameters the image, the number of upsampling times, and the model [6]. The default model used is HOG, which is the histogram of gradients.

3.2 Face Encodings

After the completion of face detection, a face encoding in the image is calculated using the face_encodings function of the face_recognition library [6]. The parameters passed to the function are the image, locations of the face in the image calculated in the previous step and the model to be used.

3.3 Face Recognition

The most important part of the system is to compare the detected faces in the images with the ones existing in the system and return the matches found. To implement this, the compare_faces function of the face_recogntion library is used that accepts the face encodings of the unknown image to be compared with the list of known encodings along with the tolerance level. The tolerance level for the proposed system is 0.45. The function calculates the distance between the two encodings by calling the face_distance function to measure how similar the two faces are.

4 Methodology

The methodology followed while building the touchless attendance system is as follows:

- The first task was designing a proper user-friendly and easy-to-use user interface. This proposed UI is on PyQt5 and Android studio.
- Subsequently, a database (MySQL) was built while keeping in mind the requirements and the developed user interface.
- Due to the multiple users present in the system, the dual login feature in the application has been added. Further, dual authentication functionality has been added for security purposes.
- That is after adding the login details, the person's face will also be recognized.
- The admins will also have to log in through the mobile application and perform a camera setup.
- The live feed from these clients (one at entry and the other at exit) will be sent to our server application that will perform the processing and send the recognized images back to the clients.
- At the backend, face recognition has been implemented with the help of dlib and face recognition library.
- Finally, the admins can access the admin dashboard for performing additional operations.

Thus, these are the steps through which the proposed application has been built and deployed.

5 System Design and Working

UpAIsthiti, a touchless attendance system which mainly works on the Ubuntu and Windows operating system. The system design of UpAIsthiti is as follows:

5.1 Users of the System

Administrator (Attendance in charge)
The administrator is in charge of keeping track of the system's users' daily attendance. The administrator can complete the following tasks (Fig. 1).

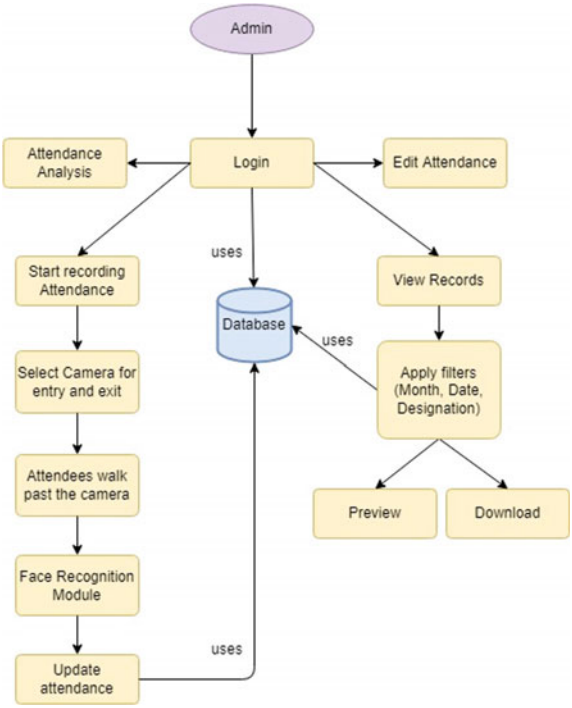


Fig. 1 Administrator module

Marking attendance

The system makes use of three devices for the whole process. One mobile phone is used to record the check-in time, and the other is used to record the check-out time. One laptop is used as a server for processing the attendance and storing in the database. The institution's unique ID is used to identify the individual. The video server is continuously open, and frames are being captured at the rate of one frame per second. Face recognition takes place and the attendance is recorded.

View records

This feature enables the administrator to view pre-saved presence. Users can use filters like select a month, date, and personal profile. Depending on the filter used, the results will be displayed in the form of a table.

Edit attendance

In a situation where the superadministrator needs to change the details of a particular user's presence, the super-administrator can search for the user with his or her unique ID and make the necessary changes like marking a person late, absent or present for a particular day.

Analysis

The analysis module present on the admin dashboard provides a visual analysis of the attendance of the organization in the form of charts and graphs.

Superadministrator side

The superadministrator is responsible for managing the system, adding new users, updating user information, and deleting the user (Fig. 2).

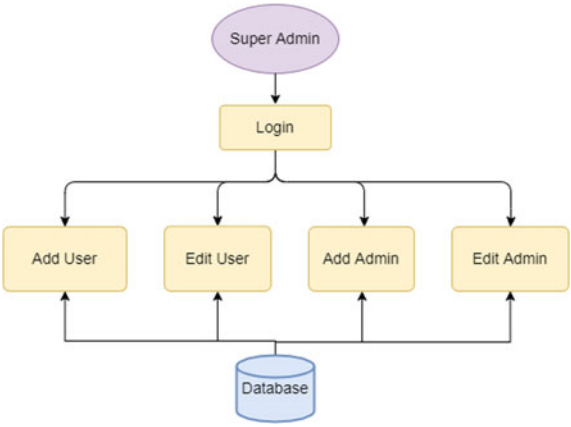


Fig. 2 Superadministrator

Add user

The superadministrator can enter the details of every user and populate the database. Once these details are submitted, encoding for the person's image is created by the system which gets stored in a dat file.

Update user

The superadministrator can update user information with this feature.

Deactivate user

To remove a current user from accessing the application, the superadministrator enters the unique employee ID of the person to be deactivated.

Add admin

The superadministrator can also add administrators, i.e., attendance in charges to the system.

Update admin

In case of any discrepancies while adding the administrator or due to any changes in the current information, the superadministrator can update the details using this feature.

Deactivate admin

To deactivate a current admin from the system, the superadministrator enters the unique employee ID of the person. Later, the access is removed accordingly.

5.2 System Implementation

Server

User interface

The user interface for the system has been developed using pyqt5. The interface is designed keeping in mind the easy accessibility of various functionalities.

Database

The database used for UpAIsthiti is developed using MySQL. There are three tables in the database: admin login, user, and attendance.

Face recognition

A set of 128 encodings is created for the image which is mapped to the user's unique ID and stored in a file for later retrieval. First, the face detection algorithm detects any human faces present in the frame. This is followed by a generation of 128-encodings for every face detected. These encodings are then matched with the encodings of known faces stored in the system and a match is found.

Receiving frames from client and marking the attendance

The server receives the individual frames sent by the client application through the socket connection. Then, the image is processed to make it fit for face recognition. After performing face recognition, the timestamp of face recognition is captured and stored for the respective camera and the attendance is marked successfully.

Sending the recognized images

After the person is identified from the image, and their attendance is marked, the processed images having the bounding box and name of the person are sent back to the client application.

Client

User interface

The user interface of the UpAIsthitiLiveFeed app has been developed on Android Studio. The live feed applications are compatible with android mobiles.
Database

The live feed apps (one for entry and one for exit) each communicate with the server and use the database of the server itself. Thus, for login purposes, the data is securely

passed to the server where the client user gets authenticated and the result is given back to the client application.

Image capturing

The Android-based client device makes use of the front camera for capturing the live camera feed and sending the frames to the server at 1 frame per second.

Recognized image retrieval

After the frames are processed by the server, they are sent back to the client to display. The server sends only the first frame in which a particular attendee is recognized to improve the bandwidth consumption and make the system more efficient.

Client–server communication
The client (mobile applications) and server (laptop) communicate by means of socket programming. The server has a static IP address to connect the client. Two server ports are needed for smooth communication. One port for the entrance part and the other for the exit part.

Security
The Android-based client consists of an authentication module wherein the attendance in charge needs to log in to the application to authenticate the identity. The login credentials are encrypted using the AES-128 encryption algorithm. The server verifies the credentials based on the entries in the database and sends a response back to the client.

6 Results and Conclusion

Thus, with the help of the face recognition library, the touchless attendance system has adequate accuracy required for proper functioning and correct results.

The Dlib model works with an accuracy of 99.38% which helps us in successful face recognition. “UpAIsthiti: A Touchless Attendance System” will thus perform as an all-inclusive platform for capturing the attendance of an organization. This system introduces an efficient and accurate method of attendance that has the potential to replace manual methods. Further, the system has no specialized hardware requirements for working. It needs only two mobiles and a laptop/computer. This automation is also efficient in the way that it will also help in eliminating fraudulent entries.

References

1. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57(2):137–154
2. Kumar A, Kaur A, Kumar M (2019) Face detection techniques: a review. Artif Intell Rev 52(2):927–948
3. Bah SM, Ming F (2020) An improved face recognition algorithm and its application in the attendance management system. Array 5:100014
4. Arsenovic M et al (2017) FaceTime—deep learning based face recognition attendance system. In: 2017 IEEE 15th international symposium on intelligent systems and informatics (SISY). IEEE
5. Kar N et al (2012) Study of implementing automated attendance system using face recognition technique. Int J Comput Commun Eng 1(2):100
6. https://face-recognition.readthedocs.io/en/latest/face_recognition.html
7. Talaviya G, Ramteke R, Shete AK (2013) Wireless fingerprint based college attendance system using Zigbee technology. Int J Eng Adv Technol (IJEAT) 2(3):201–203
8. Shukla S (2013) RFID based attendance management system. Int J Electr Comput Eng 3(6):784
9. Chew CB et al (2015) Sensors-enabled smart attendance systems using NFC and RFID technologies. Int J New Comput Archit Appl 5:19–29

Defect Discrimination of Mango Using Image Processing Techniques

Ashoka Kumar Ratha, Santi Kumari Behera, Nalini Kanta Barpanda, and Prabira Kumar Sethy

Abstract Defect discrimination is a crucial factor in evaluating the quality of mango. The flaw mark on the flesh remark about the defunctness of mango fruit. The internal defect of mango fruit is often directly proportionate to the area of flaw marks on the skin. The defect discrimination of mango fruit manually with respect to the area of flaw marks is challenging, erroneous, and time-consuming. Hence, the development of automated techniques for defect discrimination of mango fruit is needed. Here, an image processing method is developed to discriminate the defective flaw mark on the skin of mango fruit into three categories, i.e., good, average, and bad. The experimentation reveals that the proposed method successfully categorizes the mango as per the area of flaw marks on the mango fruit's skin.

Keywords Defect discrimination · Mango fruit · Flaw marks · Image processing techniques

1 Introduction

The mango (Mangifera indica Linn) is India's most popular fruit and is regarded as the "King of Fruits." The fruit is grown across the largest area, 2312 thousand ha, and produces roughly 15.03 million tons, accounting for 40.48% of global mango production [1]. India is the world's leading mango producer, accounting for over half of global mango production [2]. This mango fruit is not only tasty, but it also has a high nutritional value [3]. One of the most surprising nutritional facts about fresh mango is that it contains about 67% of the daily value (DV) for vitamin C [4]. This water-soluble vitamin supports your immune system, aids in iron absorption,

A. K. Ratha (✉) · N. K. Barpanda · P. K. Sethy
Department of Electronics, Sambalpur University, Burla, Odisha 768019, India
e-mail: ashok41ratha@gmail.com

N. K. Barpanda
e-mail: nkbarpanda@suniv.ac.in

S. K. Behera
Department of Computer Science and Engineering, VSSUT, Burla 768018, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_51

and encourages cell development and repair [5]. Mango is also high in the minerals copper and folate, essential nutrients during pregnancy because they promote healthy fetal development and growth [6]. Manual sorting and grading is a typical method of sorting and grading. However, this method has significant drawbacks, including low precision, labor-intensiveness, and subjectivity [7]. Machine vision detection systems have now made it possible to overcome these constraints accurately and non-destructively, thanks to the invention and development of image processing techniques [8]. The primary objective of this work is to design an algorithm that can identify defects of mango fruits-based flaw marks on the skin by digital image analysis. External distortions or damages are caused due to several reasons, as per "Mango Defect Guide (2014)" [9]. Some are stem end ruins, dendritic spots, anthraces, sap burn, skin cooking, and lenticels. All the above deformities are outside imperfections that influence the nature of mango skin. Although much advancement has been made in natural product programmed reviewing around the world, the recognition of organic product surface deformities is considered the most troublesome errand and has happened to the restricting components of computerized evaluating frameworks [10]. Subsequently, the present necessity of the market is for computerized frameworks that improve the procedure of value control by distinguishing different skin surface deformities that corrupt the nature of mangoes. Some researchers reported on defect detection of mango fruit using image processing and machine vision. Patel et al. [11] use monochrome cameras to construct a computer vision system for mango flaw detection and test its ability to detect the fault. Based on the preceding information, an algorithm was created, and its performance was assessed in terms of accuracy, efficiency, and average inspection time. The algorithm for flaw identification had an average accuracy and efficiency of 88.75% and 97.88%, respectively. Monochrome computer vision systems have proven to be quite effective at detecting a variety of common exterior flaws, such as a black lesion, mechanical damage, and so on [11]. Bacula et al. [12] looked into automating the detection of a mango flaw produced by cecid insects, which can affect a large amount of the manufacturing output. Multiple faults in a single mango image were detected and localized utilizing object detection frameworks based on CNN. Huang et al. [13] employed HALCON software to preprocess the original image, such as threshold segmentation and morphological denoising. Finally, the defect area is determined after the target features have been removed. Internal quality evaluation [14], grading [15], and variety identification [16] of mango fruit have also been reported in several studies.

2 Proposed Method

The Proposed image processing technique for defect identification of mango fruits using image analysis is illustrated in Fig. 1. The proposed approach involves grayscale conversion, thresholding, binarization, and morphological operation. The experimentation is conducted in HP pavilion core i5, window 10, and MATLAB 2021a platform.

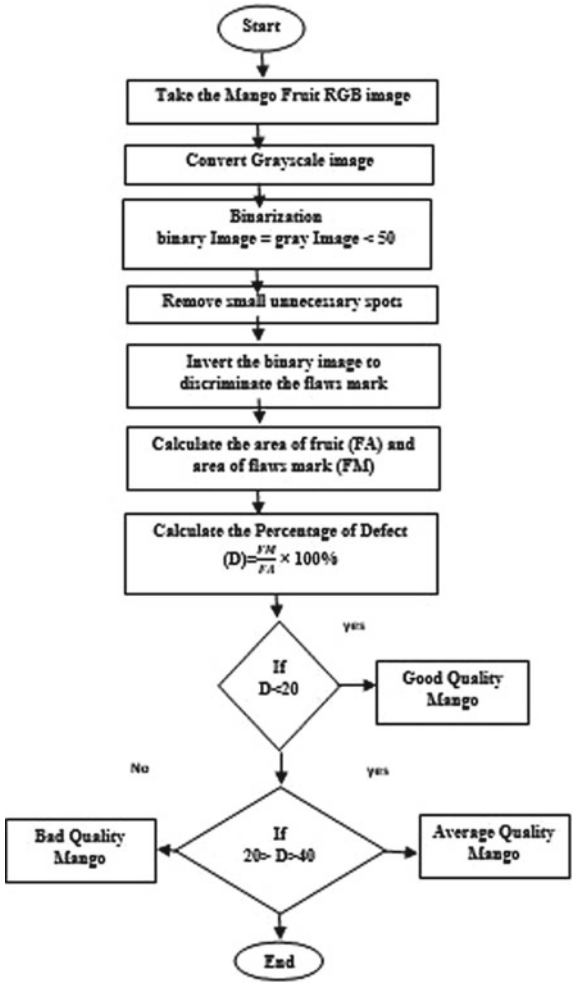


Fig. 1 The proposed method for defect discrimination of mango fruit

More than 100 images of mango fruits of different categories are taken to evaluate the methodology. A smartphone camera captures all the images with a 64 megapixel resolution with a height of 50 cm in normal daylight with white background. Special care is taken during image collection that no shadow or any other unnecessary portion is included. The images are in jpg or png format. The RGB images are then converted to grayscale images. After grayscale conversion, the binarization operation is executed. Here, a condition is applied that the resulted image only contains pixel values of less than 50. Then, clean the binarized image to remove the unwanted pixels. Finally, invert the cleaned binary image to get the flaws marks that appeared on the skin of mango fruit. The mango fruit and flaws mark area is calculated by 4-neighbor connectivity operation (*bwconncomp*) and morphological operation (*regionprops*).

After calculating areas, the percentage of defects is calculated by taking the ratio of the area of flaw marks and the fruit area by multiplying 100.

Further, the condition applied to categorize the mango fruit into three classes, i.e., good, average, and bad. Suppose the defect area is less than 20%, it is treated as good quality. If the defective percentage is between 20 and 40%, it treats as average quality, and others are bad quality mango fruit.

3 Result and Discussion

The mango fruit defect discrimination is determined by executing the various operation of image processing. Here, about 100 images of mango fruit are considered to evaluate the proposed technique. Here, for demonstration purposes, four images are illustrated with their operation steps. In the four images, one has good skin (no flaw marks), one is flaw marks with a small area, one has significant flaw marks, and another one has damaged skin (high flaw mark).

Figure 2 illustrates the different steps of mango fruit defect discrimination with no flaw mark. The binarized image is blank, resulting in the percentage of the defect being zero and resulting as good quality mango.

Figure 3 illustrates the different steps of mango fruit defect discrimination with fewer flaw marks. The binarized image contains very less area, resulting in the percentage of defect area being 16% and good quality mango.

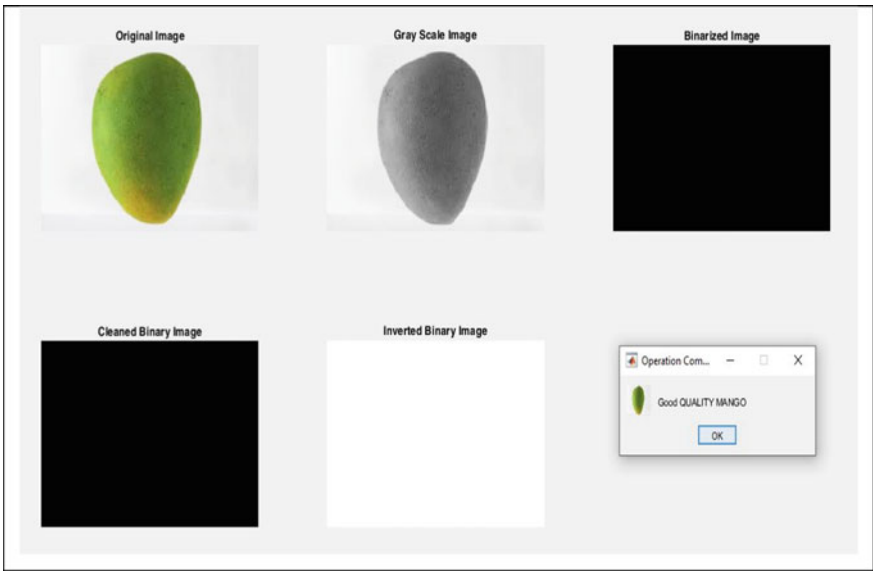


Fig. 2 Steps of defect discrimination of mango fruit with no flaw mark

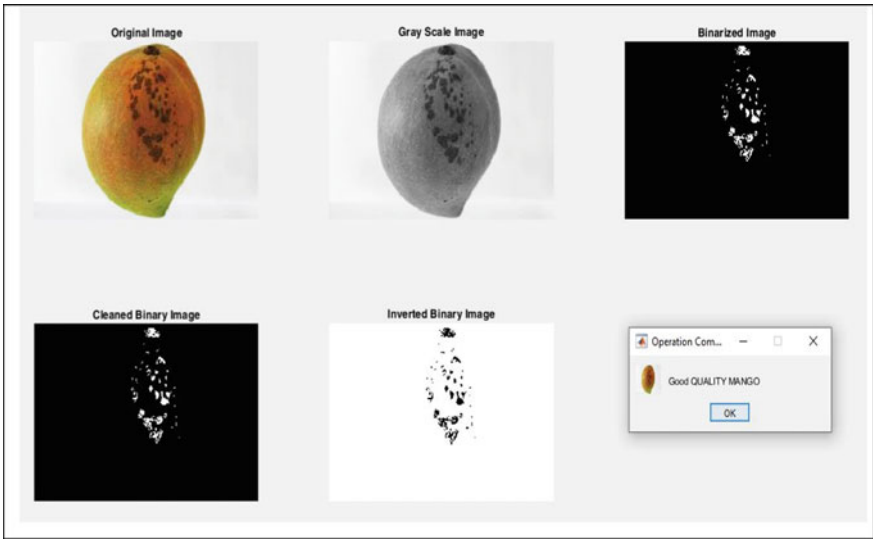


Fig. 3 Steps of defect discrimination of mango fruit with minor flaw marks

Figure 4 illustrates the different steps of mango fruit defect discrimination with significant flaw marks. The binarized image contains minimal area, resulting in a percentage of defect area of 28% and an average quality mango.

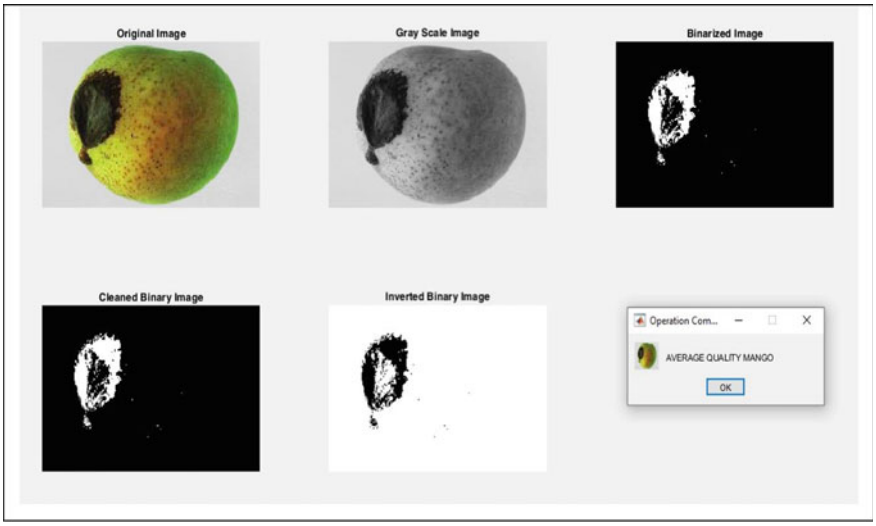


Fig. 4 Steps of defect discrimination of mango fruit with average flaw mark

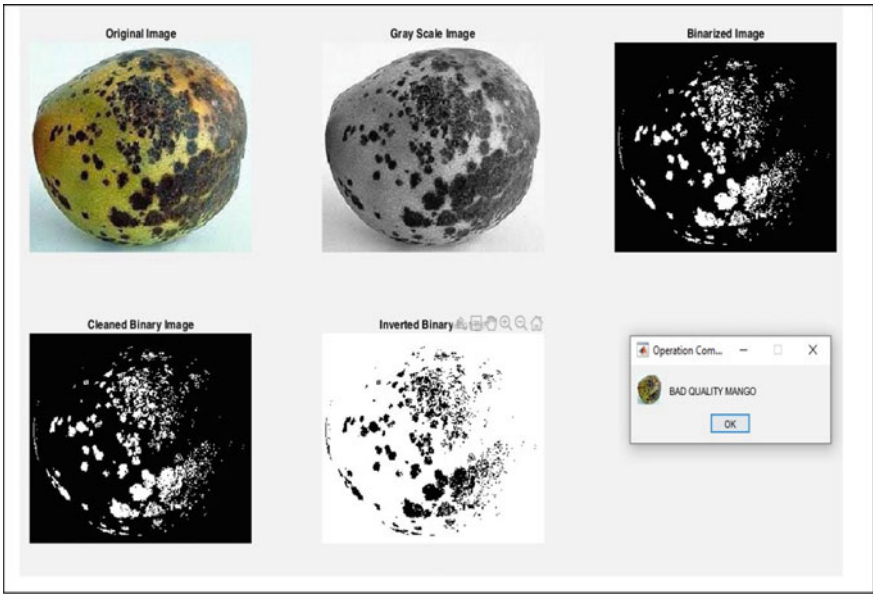


Fig. 5 Steps of defect discrimination of mango fruit with average flaw mark

Figure 5 illustrates the different steps of mango fruit defect discrimination with significant flaw marks. The binarized image contains very less area, resulting in the percentage of defect area is 48% and resulting in bad quality mango.

4 Conclusion

The mango fruit defect discrimination is an essential aspect of the fruit industry. The sorting of mango fruit as per its quality to generate the actual value of money is customer demand. This proposed technique easily detects the defected area, calculates the affected percentage, and categorizes it into proper class as per its flaws mark. This methodology is helpful for the fruit industry.

References

1. Agri Exchange-APEDA. Available: https://agriexchange.apeda.gov.in. Access: 24 April 2022
2. Mango-National Horticulture Board. Available http://nhb.gov.in/report_files/mango/mango.html. Access: 24th April 2022
3. Lebara VR, Wee YJ, Ye W, Korivi M (2021) Nutritional composition and bioactive compounds in three different parts of mango fruit. Int J Environ Res Public Health 18(2):741. https://doi.org/10.3390/ijerph18020741. PMID: 33467139; PMCID: PMC7830918

4. Ediriweera MK, Tennekoon KH, Samarakoon SR (2017) A review on ethnopharmacological applications, pharmacological activities, and bioactive compounds of Mangifera Indica (Mango). Evid Based Complement Alternat Med 2017:6949835. https://doi.org/10.1155/2017/6949835. Epub 31 December 2017. PMID: 29456572; PMCID: PMC5804368
5. Carr AC, Maggini S (2017) Vitamin C and immune function. Nutrients 9(11):1211. https://doi.org/10.3390/nu9111211. PMID: 29099763; PMCID: PMC5707683
6. Chen X, Zhang Y, Chen H, Jiang Y, Wang Y, Wang D, Li M, Dou Y, Sun X, Huang G, Yan W (2021) Association of maternal folate and vitamin B12 in early pregnancy with gestational diabetes mellitus: a prospective cohort study. Diabetes Care 44(1):217–223. https://doi.org/10.2337/dc20-1607. Epub 6 November 2020. PMID: 33158950; PMCID: PMC7783943
7. Karlsson P (2016) Is manual harvest really better than mechanical harvest? Retrieved from https://www.bkwine.com/features/winemaking-viticulture/raw-truth-manual-mechanical-harvest/
8. Chen YR, Chao K, Kim MS (2002) Machine vision technology for agricultural applications. Comput Electron Agric 36(2–3):173–191
9. Mango Defect Guide (2014). Available https://www.daf.qld.gov.au/__data/assets/pdf_file/0019/70093/Mango-Defect-Guide.pdf
10. Sharma R, Singh VP (2017) A research on automatic detection of defects in mango fruit through image processing and machine learning techniques. J Adv Sch Res Allied Educ 12:1258–1268
11. Patel KK, Kar A, Khan MA (2021) Monochrome computer vision for detecting common external defects of mango. J Food Sci Technol 58:4550–4557. https://doi.org/10.1007/s13197-020-04939-9
12. Baculo MJC, Ruiz C, Aran O (2021) Cecil fly defect detection in mangoes using object detection frameworks. In: Advances in computer graphics. CGI 2021. Lecture notes in computer science, vol 13002. Springer, Cham. https://doi.org/10.1007/978-3-030-89029-2_16
13. Huang T et al (2021) Mango surface defect detection based on HALCON. In: 2021 IEEE 5th advanced information technology, electronic and automation control conference (IAEAC), pp 2627–2631.https://doi.org/10.1109/IAEAC50856.2021.9390783
14. Raghavendra A, Guru DS, Rao MK (2021) Mango internal defect detection based on optimal wavelength selection method using NIR spectroscopy. Artif Intell Agric 5:43–51
15. Jha SN, Narsaiah K, Sharma AD, Singh M, Bansal S, Kumar R (2010) Quality parameters of mango and potential of non-destructive techniques for their measurement—a review. J Food Sci Technol 47(1):1–14
16. Pandey C, Sethy PK, Behera SK, Rajpoot SC, Pandey B, Biswas P, Panigrahi M (2021) Evaluation of transfer learning model for mango recognition. In: Intelligent manufacturing and energy sustainability. Springer, Singapore, pp 467–474

Battery Management System in EV Applications: Review, Challenges and Opportunities

R. Shanmugasundaram, C. Ganesh, P. Tamilselvi, C. S. Ravichandran, and P. S. Mayurappriyan

Abstract Rising oil prices, rapid depletion of oil reserves, and environmental concerns are driving demand and penetration of electric vehicles (EVs) in the market. One of the very important part in Electric vehicles is the Battery Management System (BMS), which ensures the EVs reliability and also reduce the battery's maintenance cost, increase the safety and protection of battery and increase the battery's life. This paper discusses the various functions of BMS, various challenges and opportunities in the development of BMS to meet the safety and future needs.

Keywords Battery · BMS · Cell balancing · Co-estimation · Electric vehicle · Lithium ion · Kalman filter · SOC · SOH

1 Introduction

The most preferred energy storage device for EVs is the Li-ion battery due to higher power and energy densities. However, for the safe and reliable operation of batteries, it requires an effective BMS in EVs. The specific requirements of BMS in EVs are Cell charge level monitoring to protect the battery from hazards by preventing deep charging as well as deep discharging under operating conditions and thereby increasing battery's life time, Cell balancing to equalize the charge in each cell [1–4] and accurate State of Charge (SOC) [5–10] and State of Health (SOH) estimation [8–11] of the battery for the accurate prediction of battery's health and remaining charge left for future use. So, research is going on to develop an effective BMS to fulfil the

R. Shanmugasundaram (✉) · C. S. Ravichandran
Sri Ramakrishna Engineering College, Coimbatore 641022, India
e-mail: rsseee@srec.ac.in

C. Ganesh
Rajalakshmi Institute of Technology, Chennai 600124, India

P. Tamilselvi
Sri Eshwar College of Engineering, Coimbatore 641202, India

P. S. Mayurappriyan
Kumaraguru College of Technology, Coimbatore 641049, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_52

abovementioned requirements and meet the safety standards of EVs. Even though, using an electric vehicle has several advantages, several factors need to be considered to improve the design to compete with existing markets, such as fast charging, fast acceleration and deceleration, no emission of harmful gases, and independence from fossil fuels. The various factor that bother the users of EVs are cost, power and energy density, cyclic life, fast charging, specific power output, safety and protection, operating environment, etc. Researchers are also working towards improving energy and power needs using super-capacitors, and control systems [12–14].

1.1 Need for BMS

As battery technology is growing, there is a need to develop an efficient BMS that can guarantee the safety of the battery subjected to internal electrochemical reaction and operating temperature, reduce the operational and replacement cost of battery, increase the lifespan of battery, increase specific power delivery, and accurate SOC and SOH estimation for accurate battery's life prediction and charge remaining for further driving. The battery pack usually comprises of a number of cells joined in series and parallel combination to deliver the energy required by the EV. It is also required to design the Thermal management system to increase battery pack's heat dissipation capacity and maintaining the optimum operating temperature (30–40 °C) to increase lifecycle of battery. Figure 1 shows the block diagram of the Battery management system (BMS). The function of BMS is to measure and monitor the key parameters of the battery pack such as charging/discharging current, terminal voltage, battery pack temperature using data acquisition system and provide this data to the online parameter identification algorithm for identification and updating of the parameters of the battery model. These updated parameters are then sent to a state observer built on the battery's state-space model to accurately estimate the battery's SOC, SOH, and State of Life (SOL). The estimation accuracy of the online identification and update method of model parameters is relatively better than the traditional method [5–8, 15]. The EV power management system needs the information SOC, SOH, and SOL of the battery to allocate power based on the customer's requirements and to achieve performance efficiency of battery.

2 Features of BMS

The key features of the BMS are listed in Table 1. For a reliable SOC estimation [5, 9], it is required to measure accurately the each cell voltage, current, and temperature of the battery pack. However, the accuracy required to measure current and voltage depends on the type of battery used. Similarly, current measurement has to be accurate as it is one of the input to SOC estimation algorithm. In coulomb counting method [5–7, 15], since charge stored in the battery is estimated by integrating the current

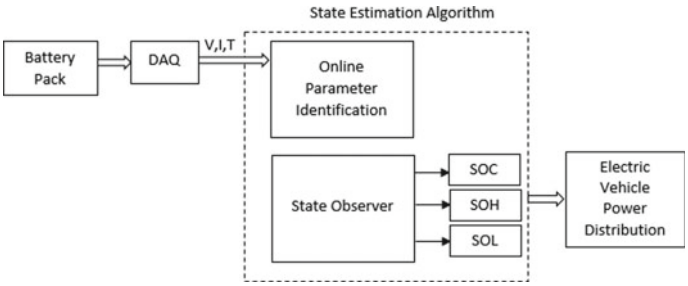


Fig. 1 Block diagram of battery management system

over time, there is requirement to use current sensor free from errors over the entire operating temperature range.

3 Techniques, Challenges and Opportunities of BMS

Several research institutions are involved in extensive research for improving the performance of BMS to satisfy the standards of EVs. The designing of BMS and its integration in EVs has many challenges and opportunities. The various research includes development of precise battery model to represent the battery's static and dynamic behaviour, development of online adaptive algorithm for accurate SOC/SOH estimation [5], modelling of cell [5– 8, 15], and state estimation for implementing active cell balancing [1–4] and improving the efficiency of the battery.

3.1 SOC Estimation Techniques

The several techniques proposed for SOC estimation are Coulomb counting, OCV approach, Electrochemical impedance spectroscopy (EIS) [7], and other online approaches [9–11]. The coulomb counting is a conventional approach, the charge stored or delivered by battery is determined by integrating current over time and this value is compared against the battery's full charging capacity. Even though it is easy to implement this method, but its limitation is measurement of initial value of SOC and accumulation of current sensor error over the period of time. In the Electrochemical impedance spectroscopy (EIS) approach, internal impedance of the battery is used for the estimation of SOC. In this technique, internal impedance is determined by passing current at different frequencies into the battery and measuring the corresponding current and voltage signals using a special signal analyzer instrument. The drawback of this method is that this process takes longer to compute the battery's internal impedance and is only apt for offline estimation. In the Open-circuit voltage

Table 1 Features of BMS

Function	Description
Cell monitoring	BMS monitors battery terminal voltage, current, and temperature
Battery model parameter identification and updating of parameter	Battery parameter that are subject to change due to change in charge/discharge rate, and battery pack temperature are identified online and updated
SOC and SOH Estimation	Model parameters are identified and updated online, and from these parameters SOC and SOH is estimated using state estimation algorithms. The SOH provide information about Remaining Useful Life (RUL) of the battery. The accurate SOC estimation helps to prevent deep discharge and over charging of the battery, and thereby prolonging the battery's life
Cell balancing	An Efficient cell balancing technology is used to balance the voltage of each cell of battery pack and control the charge/discharge rate to extend battery life
Charge control	A proper charge control strategy is needed for controlling the charge flowing either from or to the battery
Thermal management	The battery pack is designed with appropriate thermal management strategy to dissipate the heat and maintain the operating temperature of the battery cells within safe limits so as to ensure the battery's safety, optimize the performance and increase battery's lifetime
Battery safety and protection	The BMS have an effective control mechanism to prevent the battery from deep discharge, over charging, short-circuit and maintain operating temperature within safe limits and ensure safety of the battery and users. EVs application demand compliance with ISO safety standard

(OCV) approach, the open-circuit voltage of the battery is measured, and a relationship is established between SOC and OCV. The accuracy of SOC estimation can be increased by combining Coulomb counting and OCV approaches. Recently, online adaptive techniques are becoming popular. In some of these techniques, battery is modelled as a nonlinear system. The advanced observers like Sliding mode observers and Kalman filters [5, 6] are designed for accurate estimation of SOC. These methods are based on the offline identification of battery model parameters that are considered to remain constant. But it is found that the results obtained by these methods differ from batteries modelled on different environmental conditions and SOCs. It is found that battery model parameters change drastically if SOC changes from zero to 100% at fixed discharge rate and temperature. The block diagram of recently proposed

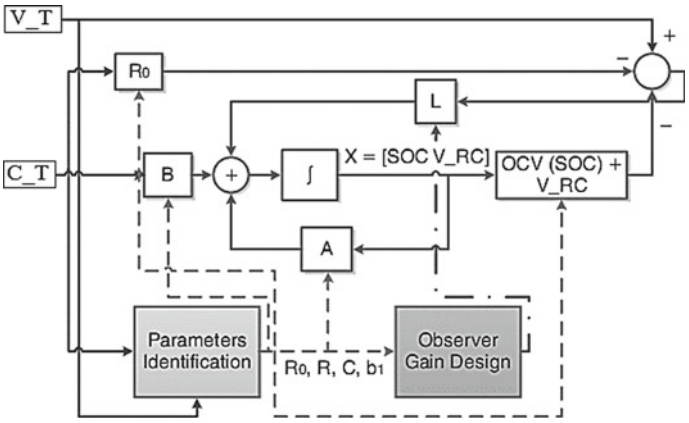


Fig. 2 Block diagram of SOC co-estimation and battery parameters

battery model parameters SOC co-estimation approach [10, 11] is shown Fig. 2. In this SOC co-estimation approach, an online parameter algorithm is used to identify the battery model parameters and with SOC-OCV piecewise linear mapped curve, the model parameters are updated continuously for precise representation of both static and dynamic behaviour of the battery. The SOC of the battery is estimated using an advanced observer which is designed based on the continuously updating battery model. The results indicates that accuracy of SOC estimation is increased by this approach and need for compensation arising due to uncertainties is eliminated.

3.2 SOH Estimation Algorithm

SOH is an indicator of how much charge is available in the battery for future use and battery charge holding capacity which indicates the battery health. Most of the literatures reports definition of SOH as ratio of actual capacity of battery that is degrading because of cycling effect and rated capacity of the battery [5–10]. The main limitation of this SOH definition is that it does not take into account battery application. Therefore, it is necessary to define SOH not only based on application but also based on age and usage history. Two main reasons for battery wear are (1) the cyclic life, and (2) the calendar life. These first aging factor is represented by increased resistance and the second by capacity degradation. Even temperature rise of the battery is an indication of battery life. For example, a 10 °C increase in temperature reduces the battery life to half. Therefore, it is challenging task to find a SOH definition for an application. Most research considers only the capacity degradation and internal resistance increase to measure SOH without considering application of the battery. The other indicators that are used predict the battery's life are End of life (EOL) and Remaining useful life (RUL). These indicators are

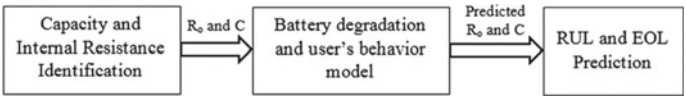


Fig. 3 Estimation of RUL and EOL in EV application

estimated by statistical analysis without considering battery application. Therefore, it is required to define SOH based on specific application demanding different energy and power at different times as well as battery's specific charging-discharging pattern. The EV behaviour is also dependent on the user's stochastic behaviour and operating environment. Therefore, an accurate battery model representing increased internal resistance and capacity degradation is needed to determine the current status of the battery by online identification of corresponding battery model parameters. The block diagram of estimation of RUL and EOL for an EV application is shown in Fig. 3. The first step in the estimation of RUL and EOL is to estimate increased internal resistance and degradation of the battery capacity in the specific EV application using an online model parameter identification algorithm. The second step is to update the battery's model by online identified values of first step and predict the future internal resistance increase and degradation of battery capacity. In final step, estimate the EOL and RUL from the data related to the user's behaviour and application specific statistical analysis tools such as Monte Carlo simulation and Bayesian analysis.

3.3 Battery Modelling

The researchers have proposed several models [5–8, 15] for the battery model which is essential for accurate SOC and SOH estimation. The battery manufacturers mainly prefer to use a model that is based on battery's electrochemical reactions to optimize the battery design. But, this type of modelling requires more memory and time for computation in order to solve the partial differential equations. Some investigators use battery current–voltage characteristics of a specific application for developing battery model. Later, battery model is developed based on statistical modelling approach, but this type of modelling does not consider the battery dynamic characteristics pertaining to operating conditions and specific application.

The limitations of above methods have made the researchers to develop models representing the dynamic behaviour of the battery. In these models, an electric circuit is used to model the battery and various battery physical phenomena is represented by the model parameters. The commonly used electrical models are (1) Thevenin-based model and (2) Impedance-based model. Frequency analysis of voltage-current characteristics of battery is used to develop the impedance-based models whereas an electric circuit with battery parameters represented by combination of resistors and capacitors is used to develop the Thevenin-based model. In the Thevenin-based models parallel R–C pairs represent the battery's relaxation's effect and the nonlinear

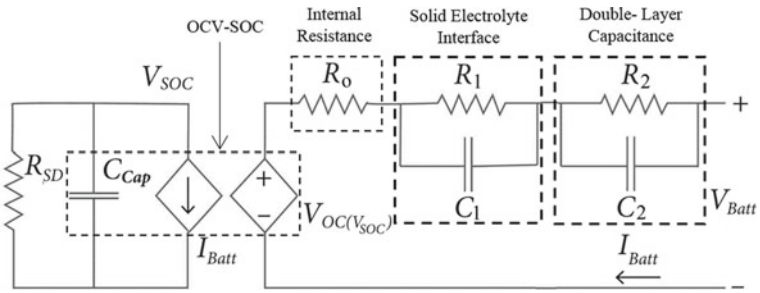


Fig. 4 Thevenin-based battery model

SOC-OCV relationship is represented by a controlled voltage source. The number of RC pairs to be used in the model is decided based on the trade-off between the model complexities and accuracy. In many literatures, it is shown that one RC pair is enough to deal with relaxation effects and improve performance. The Thevenin-based model [6] is shown in Fig. 4. In this model, internal resistance R_o represent the opposition to movement of ions in the electrolyte, parallel combination of R_1-C_1 represents short-term relaxation effect produced at the anode by the electrolyte interface, parallel combination of R_2-C_2 represents long-term relaxation effect produced by double-layer capacitance formed at cathode and anode, and combination of SOC and Q represent the OCV-SOC relationship considering hysteresis effect. The parameters of this model are likely to change when there is a change in the battery temperature, discharging-charging current rate, ageing effect, and SOC. This model can be used as an adaptive model, by identifying and updating of parameters online using online parameter identifying algorithm.

3.4 Cell Balancing

The various approaches on cell equalization [1–4] reported in the literatures are listed in Table 1. Methods of cell equalization are broadly divided into passive and active cell equalization. In passive cell equalization, more energy in highly charged cells is passed through resistor and dissipated as heat. As energy is wasted, this method is considered ineffective. With active cell equalization, more energy in highly charged cells is transferred to other cells, and these methods are considered to be more efficient. Each active cell balancing have a different level of circuit complexity, efficiency and a way of energy distribution amongst the cells. Therefore, the choice of active cell approach should be based on a trade-off between circuit complexity and efficiency that can be achieved. The most promising active cell balancing approaches are found to be cell-to-cell and module-to-cell balancing methods. In these techniques, energy in the highly charged cells is redistributed to the less highly charged cells in a module using a single dc-dc converter. In hardware, the Module Management Unit

(MMU) contains charge equalizing circuit, the PMU monitor and control the entire cell equalization process and the amount of charge per cell of the battery pack. The efficiency of up to 90% can be achieved by the cell equalizer at the module level.

4 Conclusion

The BMS has become an essential component of EVs industry. Only with proper BMS and battery technology that can provide higher energy and power density at reduced cost, it is possible to efficiently control the parameters of the battery, and ensure the safety and reliability of the battery. Therefore, research and development is needed to build an accurate battery model representing increased internal resistance and degradation of battery's capacity, develop an SOC accurate measurement algorithm for online detection and updating of model parameters, redefine SOH based on specific application, user's stochastic behaviour and operating environment to accurately predict EOL and RUL of battery, and develop simple, inexpensive, minimal control complexity and fast-acting methods to equalize all cells in battery pack module by an effective active cell balancing method.

References

1. Vulligaddala VB et al (2020) A 7-cell, stackable, li-ion monitoring and active/passive balancing IC with in-built cell balancing switches for electric and hybrid vehicles. IEEE Trans Industr Inf 16(5):3335–3344
2. Lee S, Kim M, Baek JW, Kang DW, Jung J (2020) Enhanced switching pattern to improve cell balancing performance in active cell balancing circuit using multi-winding transformer. IEEE Access 8:149544–149554
3. Haq IN et al (2014) Development of battery management system for cell monitoring and protection. In: 2014 international conference on electrical engineering and computer science (ICEECS). Kuta, pp 203–208
4. Rahimi-Eichi H, Ojha U, Baronti F, Chow M-Y (2013) Battery management system: an overview of its application in the smart grid and electric vehicles. IEEE Ind Electron Mag 7(2):4–16
5. Song Y, Liu W, Li H, Zhou Y, Huang Z, Jiang F (2017) Robust and accurate state-of-charge estimation for lithium-ion batteries using generalized extended state observer. In: Proceedings of the 2017 IEEE international conference on systems, man, and cybernetics (SMC), Banff, AB, Canada, pp 2146–2151
6. He H, Xiong R, Zhang X, Sun F, Fan J (2011) State-of-charge estimation of the lithium-ion battery using an adaptive extended Kalman filter based on an improved Thevenin model. IEEE Trans Veh Technol 60(4):1461–1469
7. Di Domenico D, Fiengo G, Stefanopoulou A (2008) Lithium-ion battery state of charge estimation with a Kalman filter based on an electrochemical model. In: IEEE international conference on control applications, San Antonio, TX, USA, pp 702–707
8. Tan X, Tan Y, Zhan D, Yu Z, Fan Y, Qiu J, Li J (2020) Real-time state-of-health estimation of lithium-ion batteries based on the equivalent internal resistance. IEEE Access 8:56811–56822
9. Zhang C, Wang LY, Li X, Chen W, Yin GG, Jiang J (2015) Robust and adaptive estimation of state of charge for lithium-ion batteries. IEEE Trans Indust Electron 62(8):4948–4957

10. Rahimi-Eichi H, Baronti F, Chow M (2014) Online adaptive parameter identification and state-of-charge coestimation for lithium-polymer battery cells. IEEE Trans Industr Electron 61(4):2053–2061
11. Shen P, Ouyang M, Lu L, Li J, Feng X (2018) The co-estimation of state of charge, state of health, and state of function for lithium-ion batteries in electric vehicles. IEEE Trans Veh Technol 67(1):92–103
12. Shanmugasundaram R, Ganesh C, Singaravelan A, Gunapriya B, Adhavan B (2022) High-performance ANFIS-based controller for BLDC motor drive. In: Smart innovation, systems and technologies, vol 243, pp 435–449
13. Shanmugasundaram R, Ganesh C, Singaravelan A (2020) ANN-based controllers for improved performance of BLDC motor drives. In: LNEE, vol 665. Springer, Singapore, pp 73–87
14. Shanmugasundram R, Zakariah KM, Yadaiah N (2014) Implementation and performance analysis of digital controllers for brushless DC motor drives. IEEE/ASME Trans Mechatron 19(1):213–224
15. Giordano G, Klass V, Behm M, Lindbergh G, Sjöberg J (2018) Model-based lithium-ion battery resistance estimation from electric vehicle operating data. IEEE Trans Veh Technol 67(5):3720–3728

Analysing Impact of Online Advertisement Using Internet on Consumer Buying Behaviour with Special Reference to Mobile Phone During Covid-19 and Pre-covid-19 Period

Jihan Mehra and Mahendra Parihar

Abstract The new coronavirus, which produces a highly contagious sickness, enters the picture. Globally, the coronavirus disease (COVID-19) has boosted the use of Internet commerce. It has resulted in an increase in the number of FTUs, or first-time e-commerce users, in India, who were previously unable to purchase online. Customers are depending on Amazon more than ever before in their social isolation and self-quarantine attempts, according to Amazon, one of the country's and the world's major e-commerce businesses (https://retail.economictimes.indiatimes.com/re-tales/impact-of-covid-19-on-onlineshopping-in-india/4115). The goal of this study is to determine the elements that influence customers' online purchases of mobile phones, as well as how that practice has altered since the global COVID-19 outbreak. An online survey was done, and data were gathered from primary sources to interpret the objectives. The goal of the study is to figure out how successful online advertising is at raising awareness and what the link is between Internet advertising and purchasing decisions.

Keywords Covid-19 · E-commerce · Marketing · Consumer choice · Internet advertising

1 Introduction

The unorganised retail sector, which comprises around 13.8 million traditional family-run neighbourhood establishments, and the organised retail sector, which accounts for less than 10% of the Indian retail industry, are vastly different. All organised brick-and-mortar establishments and Internet shopping sites fall within the organised sector. Despite the growth of India's B2C e-commerce business, the majority of Indians still prefer to purchase at their local brick-and-mortar stores because they like to touch and feel things and negotiate discounts over the counter before purchasing. The bulk of B2C e-commerce businesses in India entice customers

J. Mehra (✉) · M. Parihar
Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Mumbai, India
e-mail: vrnmhr50@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_53

to purchase online by giving incentives such as free delivery, discounts, buy-one-get-one-free, and exchange deals. However, many Indian customers, who are known for being cost-conscious and conservative as part of their value system, are unlikely to be swayed by deals and marketing. Furthermore, Internet consumers frequently encounter issues with product delivery schedules and customer care services. Customers' fear of online Web sites is heightened as a result of some e-tailers' inadequate IT infrastructure, which has resulted in personal information being hacked [7].

The advertising function serves as the centre of gravity for all other functions. Goods and services are created to meet the needs of consumers, not for the benefit of their owners. As it is often said, business does not exist in a bubble; it must interact with many elements, and advertising is no exception. Because of its unique traits of flexibility, interaction and personalisation, the Internet has developed immensely in terms of both applications and users. It has proven to be an extremely valuable instrument for communication, entertainment, education and electronic commerce. Information technology's revolutionary shift has a significant influence on people's daily life. It has changed the way we conduct business by allowing businesses to provide an endless selection of items and services to all consumers worldwide at any time. The Internet has become a popular platform for advertising. Many businesses have resorted to the Internet to sell their goods and services, and the Internet is often regarded as the world's most important direct marketing channel. Companies are putting tens of billions into online advertising in order to get a better return on their ad expenditure. Consumers now have greater discretion over how they obtain information about products and services due to the Internet. There are various elements that influence customers' desire for online content. Consumers choose when, when, what and how much commercial material they want to see. Consumers may now access an endless selection of products and services from firms all over the world via the Internet, which has lowered the amount of time and effort they spend shopping [1].

Consumers look for information online with a specific aim in mind, and that objective might impact individual behaviour and responses to online information and marketing. With the fast progress of computer technology, many businesses have used the Internet in their advertising media mix to take advantage of online capabilities. Because marketers discovered that the Internet offers more flexibility and control over advertising materials, it has become a popular advertising platform. Because the Internet may be used to effectively communicate marketing messages, both academics and practitioners are interested in learning how to maximise the value of this communication channel [2].

To completely comprehend the definition of advertising, we must first comprehend the numerous forms of advertising, which include the following:

- Through media, which can be aired media (broadcast TV and radio) or print media (newspapers, journals or flyers), information about the product can be distributed to a dispersed audience at the same time and successfully.

- It may also be done by placing the goods on display in a moving vehicle, a business or a facility. It is frequently done in an eye-catching manner to attract people's attention.
- Another method is to use billboards, which are enormous signs or boards installed along busy streets to educate people of a current products amongst all passers-by.

Furthermore, digital advertising is the process of using the Internet to convey marketing messages to a specific and targeted audience. It is beneficial for increasing website traffic and brand awareness, but digital advertising is first and foremost intended to encourage the targeted client to do a certain action, such as completing a purchase [2].

2 Literature Review

Online retailers have a number of benefits over traditional businesses: They are handy and time-saving, and they eliminate the need for travel and line waiting. They are open 24 h a day, 7 days a week, and may be reached at any time. Customers may get free and detailed information on products and services at these businesses. They also have several online tools to assist clients in comparing and purchasing different items and services. Interactivity is the primary differentiator between marketing communication on the Internet and conventional mass media, according to the study. Because the Internet provides more interactivity between customers and product/service providers, as well as better availability of information about items and services, online consumers now have more control and negotiating power than consumers in traditional storefronts. The Internet altered the power balance in customers' favour by making it relatively easy for them to compare and assess products without being harassed by salesmen. Both customers and suppliers benefit from online businesses because they cut transaction costs. However, as compared to brick-and-mortar establishments, Internet stores have several drawbacks. Customers searching for and purchasing things at online retailers are unable to have any sense of the product they view on the Internet (seeing, touching, tasting, smelling and hearing). Because there is no face-to-face connection in online retailers, customers may acquire a low level of trust and perceive a high level of danger. Although, with the use of certain software solutions available on the market, this difficulty can be lessened [3].

Furthermore, one of the research found that when a person's mind is engaged with a product, advertising has a good effect on his or her desire to look for information; thus, the first thing the person does is go to the Internet. As a result, it is advised that enterprises selling online items develop marketing strategies that convey facts about the product to consumers at a broad level of the virtual world of information.

The Internet is becoming into a new means to shop for various items and services. Even still, it is a desirable scenario for everyone to be able to touch the things that they wish to purchase. However, the Internet is now playing a larger role in making buying easier than ever before. Shopping has become lot easier thanks to the Internet, and it is

now only a click away. "Online Shopping" is a new phrase that has been established. Consumers may buy products or services directly from merchants without having to deal with any middlemen. The Internet is becoming an important marketing route, similar to previous direct marketing channels such as television and catalogues. The Internet allows consumers and merchants to communicate in both directions. The Internet offers an interactive shopping channel that is not limited by location and time [4].

With acculturation, customer buying choices are fast altering and migrating towards high-end technological items. Because of changing lifestyles and growing economic levels, products that were formerly considered luxury things have become necessities [5]. The demand for high-end items such as televisions, washing machines, refrigerators, and air conditioners has expanded significantly as disposable incomes have climbed. It is also made easier by the widespread availability of credit and the frequency of nuclear households. Increased demand for consumer durables in the market, as well as price reductions, as Indian consumers continue to place a premium on value for money.

The research investigates the influence of advertising on consumer purchasing behaviour in the electronics sector (consumer electronic items) in Maharashtra's metro cities. Because of the fierce rivalry in the market, every firm in India takes a very aggressive approach to selling their products and services through efficient promotional channels. Advertising is essential for any business since it allows them to showcase their products and services whilst also influencing customer purchasing decisions. With the advancement of technology, businesses have the ability to develop more effective advertising, which has become critical for them to maintain a competitive edge [2].

Advertising is a type of commercial communication in which a product, service or concept is pushed or sold by sending a non-personal, overtly sponsored message. Advertising sponsors are frequently used by businesses who wish to sell their products or services. The bulk of advertising is directed at groups rather than individuals, and it is frequently broadcast on television, social media and newspapers. Advertising is a very old kind of marketing with roots that can be traced all the way back to the beginning of time. Advertising allows you to reach out to a specific target in a personal and effective way. Customers are reminded of the benefits of the company's products and services through advertising [1].

3 Objective of the Study

- To study the impact of online advertisement on a consumer buying mobile phone.
- To comprehend the power of persuasion in advertising in order to purchase a thing.
- To determine the most popular advertising medium amongst the general public.
- To study the paradigm shift in the mode of purchase pre-Covid-19 & during Covid-19.

Methodology: Methodology adopted for the study includes structural approach which includes the following:

Sampling-

Google form was made and was distributed covering the topic related questions via Email and WhatsApp. The total samples collected were 250 covering the age group of 18–40 & above.

Tools used for sampling

Primary data were used in this research. A structured questionnaire was used to obtain primary data. Secondary data were gathered from many sources such as the Internet, journals, publications, and books. We employed simple statistical techniques and some simple and relevant data supplied in tabular form for analysis and interpretations. The hypothesis was tested using the chi-square test.

Hypothesis

I. H_0: The influence of Internet advertising on a customer purchasing a mobile phone is negligible.

   H_a: Online advertisements have a huge influence on a consumer purchasing a mobile phone.

II. H_0: There is no significant association in the mode of purchasing mobile phone pre-Covid-19 and during Covid-19.

   H_a: There is significant association in the mode of purchasing mobile phone pre-Covid-19 and during Covid-19.

4 Result Analysis and Findings of the Study

Consumer behaviour may be defined as the decisions and activities that impact a consumer's purchasing behaviour. What motivates customers to pick one product over another is an issue that marketers frequently examine and research. The majority of the purchase decision-making process is based on emotions and logic.

Consumer behaviour research not only aids in understanding the past, but also in forecasting the future. To have a thorough grasp of consumer purchasing patterns, the following aspects relevant to people's habits, attitudes, and priorities must be given proper weight.

1. Marketing Campaigns:

Advertisement has a stronger impact on customer purchase decisions. They have even been known to cause significant shifts in market share in competitive sectors by influencing customer purchase decisions. Regular marketing efforts can affect a consumer's purchase choice to the point that they choose one brand over another or engage in extravagant or frivolous spending. If conducted at regular intervals, marketing efforts can even remind customers to purchase for less attractive things such as health products or insurance plans [6].

2. Economic Conditions:

Consumer spending decisions are recognised to be heavily impacted by the market's economic status. Consumers are considered to be more confident and eager to make purchases in a strong economic situation, regardless of their own financial obligations [6].

3. Personal Preferences:

Consumer behaviour is impacted by many shades of likes, dislikes, priorities, morals and values on a personal level. In some dynamic industries, such as fashion, cuisine and personal care, the consumer's personal vision and opinion on style and pleasure may become the most powerful influencing force. Though advertisements can affect these criteria to some level, a consumer's personal preferences and dislikes have a higher impact on the final purchase they make [6].

4. Group Influence:

Consumer decisions are also observed to be influenced by group influence. The primary influential group, which includes family members, classmates, and near relatives, as well as the secondary influential group, which includes neighbours and acquaintances, are thought to have a higher impact on a consumer's purchasing decisions [6].

5. Purchasing Power:

The purchasing power of a customer has a significant impact on consumer behaviour. Before acquiring any items or services, buyers usually assess their purchasing power. The product may be wonderful, but if it does not meet the purchasing power of the consumer, it will have a significant influence on its sales. Segmenting customers based on their purchasing power will aid in identifying qualified customers and achieving better outcomes.

Understanding, analysing and tracking customer behaviour is vital for a marketing department to maintain a successful position in the marketplace. Apart from the five elements described above, there are several more factors that impact customer behaviour [6] (Table 1).

A total of 250 people responded to the poll, with 210 of them being between the ages of 18 and 28, 14 being between the ages of 29 and 39 and the remaining 26 being between the ages of 40 and 50. (40 years and above). The above data also show that the age group advertising medium (18–28 years) had the largest percentage of responders (84%) (Table 2).

Table 1 Age group of the respondents

Age (in years)	Number of respondents	Percentage (%)
18–28	210	84
29–39	14	5.6
40 & above	26	10.4
Total	250	100

Source Our study

Table 2 Gender of the respondents

Male	Female	Prefer not to say	Total
164	84	2	250

Source Our study

Out of the total 250 respondents, there are 65.6% male, 33.6% female and 0.8% preferred not to say (Fig. 1).

In reference to Fig. 2, 78.4%, i.e. 196 respondents preferred social media as their preferred source of advertisement; 16%, i.e. 40 respondents preferred television advertisement; 2.8%, i.e. 7 respondents most preferred advertisement medium is radio; 2%, i.e. 5 respondents prefer email, and 0.8%, i.e 2 preferred email advertisement. Hence, from the above information, it can easily be inferred that youth and even the middle age group prefer social media to be their source of advertisement.

According to the above graph, 41.6% of respondents (104 respondents) believes advertisement input stays in their minds for a few hours; 21.2% (53 respondents) believes advertisement input stays in their minds for a week; 20% (50 respondents) thinks advertisement input lasts in their minds for a day; 10% (25 respondents) thinks advertisement input will remain in their minds for more than a month, and the

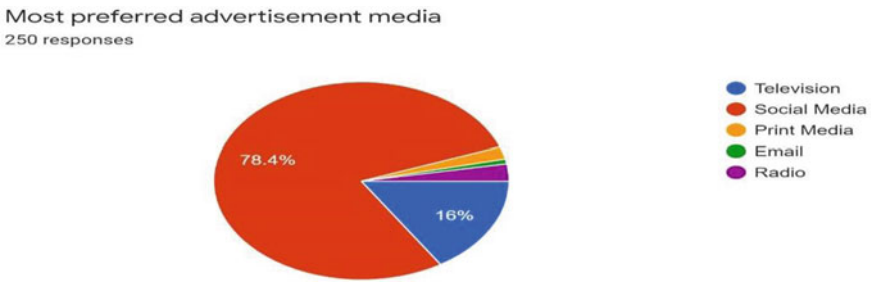


Fig. 1 Respondents view towards most preferred advertisement media. *Source* Our study

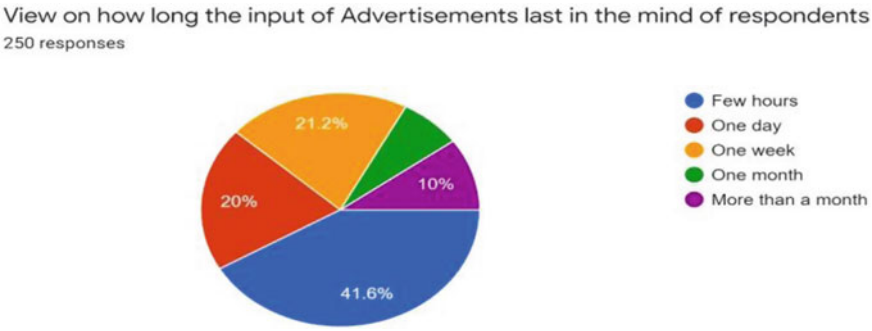


Fig. 2 View on how long the input of advertisements last in the minds of respondents. *Source* Our study

remaining 7.2% (18 respondents) believes advertisement input will remain in their minds for more than a month.

This Fig. 3 that 55.2% of respondents is influenced to purchase the mobile because they are curious about it; 41.2% is influenced by suggestions about the mobile; 40% is influenced by the interesting advertisements; 23.2% is influenced because the price of the mobile was low; 5.6% respondents are influenced to purchase the mobile because of their intention to imitate the actors in the advertisement (Fig. 4).

According to the survey conducted, 79.6%, i.e. 199 respondents preferred Amazon, whilst 12.8%, i.e. 32 preferred Flipkart, and the rest 7.6%, i.e. 19 preferred other e-commerce sites for their purchase of mobile phones.

According to Fig. 5, 30.8%, i.e. 77 preferred OnePlus; 28.4%, i.e. 71 preferred Apple, 21.2%, i.e. 53 preferred Samsung; 6%, i.e. 15 preferred Xiaomi; 2.8%, i.e. 7 preferred Motorola; 2.4%, i.e. 6 preferred Vivo; 1.6%, i.e. 4 preferred Oppo, whilst the rest 6.8%, i.e. 17 preferred other brands. Hence, this data clearly depict that OnePlus, Apple and Samsung are the most preferred mobile phone brands these days. The

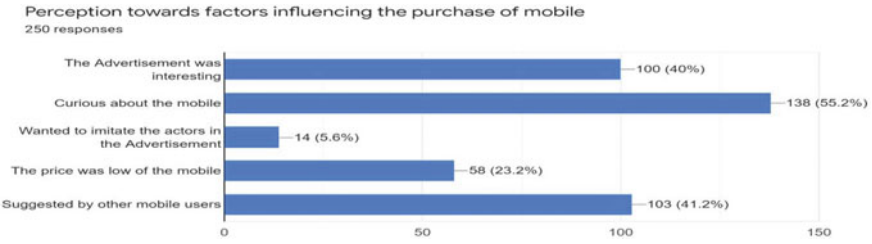


Fig. 3 Perception towards factor influencing the purchase of mobile. *Source* Our study

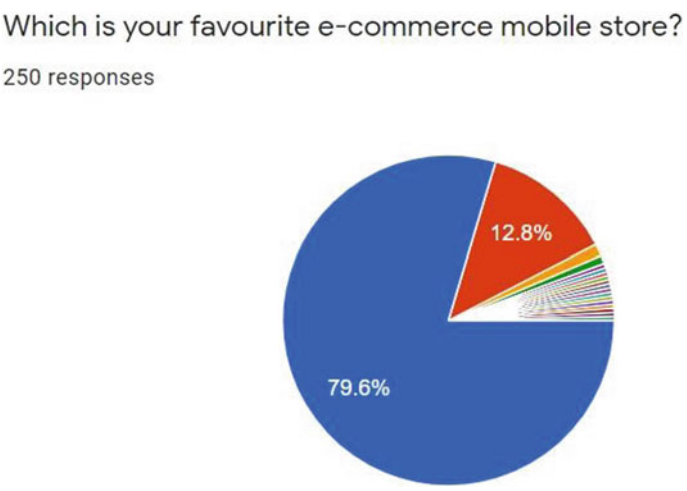


Fig. 4 Favourite e-commerce mobile store. *Source* Our study

above figures clearly depict that there have been favourable number of respondents who have changed their mode of purchase from offline to online during the Covid-19 pandemic. But, to assure whether the change was significant, a chi-square analysis was performed, and two hypothesis were assumed (Figs. 6 and 7).

Testing of Hypothesis 1:

H_0: There is no significant association in the mode of purchasing mobile phone pre-Covid-19 and during Covid-19.
H_a: There is significant association in the mode of purchasing mobile phone pre-Covid-19 and during Covid-19.

The result of the chi-square analysis is shown in Table 3.

Since chi-square value (99.369) is more than critical value (3.841), H_0 is rejected. The chi-square analysis thus reveals that there is significant association in the mode of purchasing mobile phone pre-Covid-19 and during Covid-19.

Figure 8 clearly depicts that 39.6% respondents strongly agree to the significant role of online advertisement; 36.8% agrees to the significant role of online advertisement; 16.8% was neutral about the influence of online advertisement; 4.4% disagrees to the view that online advertisement has significant influence on the consumer behaviour, and the remaining 2.4% strongly disagrees to the view that online advertisement has significant influence on the consumer behaviour.

Testing of Hypothesis 2:

H_0: The influence of Internet advertising on a customer purchasing a mobile phone is negligible.
H_a: The influence of Internet advertising on a customer purchasing a mobile phone is enormous (Table 4).

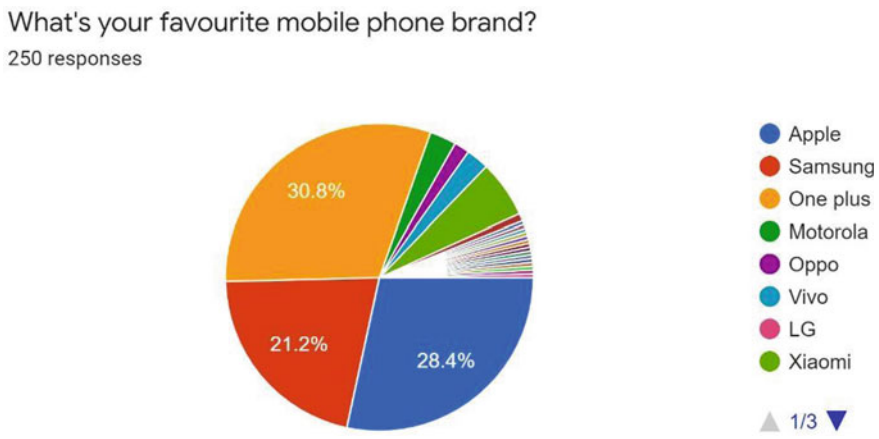


Fig. 5 Most preferred mobile brand. *Source* Our study

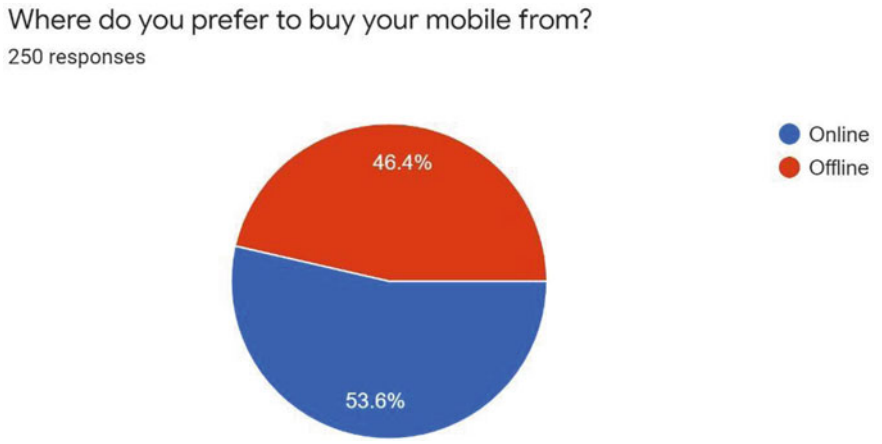


Fig. 6 Mode of purchase pre-Covid-19. *Source* Our study

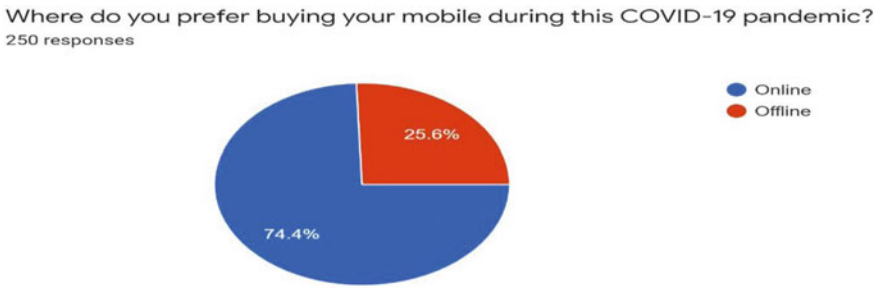


Fig. 7 Mode of purchase during Covid-19. *Source* Our study

Table 3 Testing of 1st hypothesis

Chi-square value	Critical Value	Significance
99.369	3.841	Highly significant

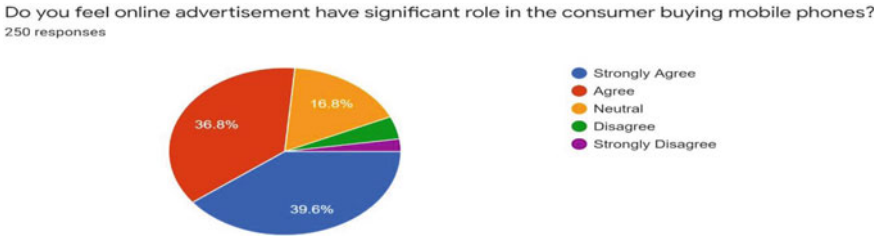


Fig. 8 View of respondents on convincing power of online advertisements towards purchasing of new mobile. *Source* Our study

Table 4 Testing of 2nd hypothesis

Chi-square value	Critical value	Significance
153.72	9.49	Highly significant

Since chi-square value (153.72) is more than table value (9.49), H_0 is rejected. The chi-square analysis thus reveals that there is significant impact of online advertisement on a consumer buying mobile phone.

5 Conclusion

Advertisement serves as a motivation for people to buy a company's products or services. Businesses utilise advertisements as one of their primary techniques for promoting and disseminating information about their products and services. Advertisements must not be deceptive since this may have a negative influence on customers. Effective advertising also aids in the growth of a company's sales. The goal of this study was to look at the influence of Internet marketing on customer purchasing behaviour and how that changed before and during Covid-19. According to the findings, Internet advertising is the respondents' favourite type of advertisement, and it has a major influence on their purchasing behaviour. Also, the respondents have shifted from offline to online mode of purchasing during Covid-19 pandemic. This research paper may help you learn about the most popular advertising medium, how commercials impact shopping patterns, the most memorable parts of advertisements and much more.

References

1. Sathya P, Indirajith R (2016) A study on purchase behavior of consumer durable goods with special reference to Tiruvarur District
2. Deshpande M, Rokade Y, Darda PS (2019) The impact of advertisement on consumer buying behavior in electronic industry
3. Harfoushi O, Alfawwaz B, Obeidat B, Obiedat R, Faris H (2013) Impact of internet advertisement and its features on e-commerce retail sales: evidence from Europe
4. Ghafoor A, Muhammad Ismail A, Fauzia C, Ghazal MF (2015) Effects of advertisement on consumer's buying behaviour with references to FMCGs in Southern Punjab-Pakistan
5. Abayia M, Khoshtinat B (2016) Study of the impact of advertising on online shopping tendency for airline tickets by considering motivational factors and emotional factors

Web Resources

6. https://www.marketing91.com/5-steps-consumer-buying-behaviour/
7. https://retail.economictimes.indiatimes.com/re-tales/impact-of-covid-19-on-onlineshopping-in-india/4115

Comparative Analysis of Intra- and Inter-Prediction Compression Techniques for Endoscopic Videos

Suvarna Nandyal and Heena Kouser Gogi

Abstract The cloud-based healthcare system has opened up new possibilities in the medical profession. By monitoring the whole digestive tract, endoscopic technology has improved the diagnosis of gastrointestinal illnesses and diseases. Image compression enhances the frame rate, which helps the diagnostic procedure. The aims of presenting this paper are to put light on brief review of image compression into two categories intra- and inter-prediction for endoscopy video compression, which is efficient and simple. Our research work shows an acceptable compression performance; the initial phase of work is an implementation of intra-prediction, and the second phase of work is inter-prediction technique, which shows a mark able performance in term of PSNR value, bit rate and compression ratio.

Keywords PSNR · Intra-prediction · Inter-prediction · Bit rate · Compression ratio

1 Introduction

The goal of our research is to improve video compression algorithms that can be needed to secure endoscopic video space in cloud-based healthcare systems. Video endoscopy is a minimally invasive operation that has been found to greatly improve diagnosis. In this work, we offer our answers using two methods: intra-prediction and inter-prediction, as well as a comparison of the two. During an endoscopic procedure, a long, flexible tube (endoscope) is introduced through the mouth and into the oesophagus. Both diagnostic and therapeutic aims are possible with flexible endoscopy. The surgeon is obligated to examine the lesions and organs that must be surgically treated, as well as to undertake any postoperative examinations that are required. In order to provide patients with more effective therapeutic options,

Patent ID: 202041040000A

S. Nandyal
Department of CSE, PDA College of Engineering, Kalaburgi, Karnataka, India

H. K. Gogi (✉)
Department of ISE, Research Scholar PDACE, Kalaburgi, Karnataka, India
e-mail: heena.shakir@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_54

the surgeon must also analyse and select innovative intervention approaches for the treatment of digestive disorders [1]. As a result of the procedure, the surgeon utilises the video-recorded endoscopic session for medical investigation, diagnosis, and therapy. Let us know about imaging in the gastrointestinal tract, which requires a few essential steps (Fig. 1). Fibre optic light transmission is the source for the light, magnification via a lens system, conversion of reflected photons to a signal via a charged coupled transfer device, signal reconstruction, and projection onto a monitor. To record, preserve, print, and communicate these pictures, high-speed microcomputers with digital image capture (frame grabbers) and network boards coupled to video processors are required [2]. Let's take a look at the endoscopic storage needs at the hospital. Consider the following scenario: A hospital has eight operating rooms, each of which performs three to two operations every day, three of which are endoscopies [3]. A single hospital recording all surgical procedures generates 189.8 GBytes of video data per day and 67.7 TBytes per year for surgery, 7.8 Tbytes of endoscopic data per day and 949 Tbytes per year for endoscopy data, according to our estimations. This corresponds to 62.9 PBytes each year for Poland as a whole. If these recordings are taken into account, they must be kept for at least 20 years following the patient's last visit, as required by law. In terms of storage, a hospital's IT infrastructure must contend with a hurdle. The above calculation assumes that you will only be able to capture one full HD video. Medical equipment with 4 K or even 8 K resolutions (such as endoscopes) is presently available on the market. As a result of the employment of these devices, the volume of data generated will increase by a factor of four or sixteen. New video encoding techniques that allow for more compression whilst keeping medical-grade quality are needed to compensate. As a result, compression has become a critical demand in the medical industry for both storage and video quality.

There are two types of video compression methods now in use:

1. Lossy compression
2. Lossless compression.

Lossy compression: This type of compression has compression ratios ranging from 10× to more than 100×, and it is usually utilised for final video dissemination. Lossy compression achieves high compression at the expense of visual information, and it often involves a large computing effort on the encoder's part. With the most modern video codec, however, the encoder may control the compression ratio and the resulting visual quality for a variety of application. We can obtain video quality that is visually indistinguishable from the source for telemedicine. Active compression video coding technologies include H.264/AVC and H.265/HEVC.

Lossless compression: Data compression that is lossless is used to transfer or store data whilst keeping the same quality as the original. Strong compression ratios of 2× to 6× are common; however, encoder and decoder implementations employ significantly fewer computer resources. JPEG-2000, JPEG-XS, and VC-2 are amongst the codecs in this category [4].

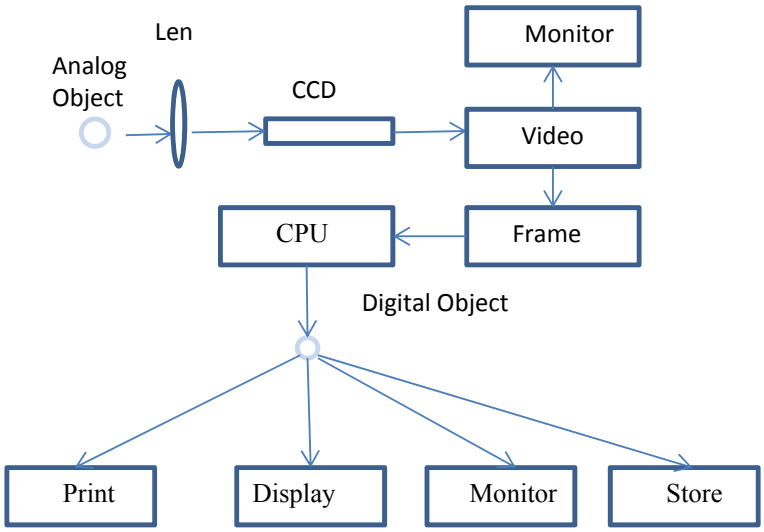


Fig. 1 Block diagram of endoscopic digital image management system

2 Related Work

Let's have a look at the compression algorithms for high-definition video coding and rate-distortion optimization, which employ an inter–intra prediction. Due to motion correction, inter-frames are extensively compressed, and any embedding message within them can dramatically reduce compression efficiency. Inter-frame payload is also assumed to be lower than intra-frame payload due to the smaller amount of leftover data. Texture masking is employed in inter-frames. We can see that the data are hidden throughout the compression process, and the proposed approach takes the reconstruction loop into consideration [5]. Our research implementation deals inter–intra prediction distinctly to achieve the compression parameters. Intra-frame coding is a data compression technique that reduces file sizes and bit rates without compromising quality. Because close pixels in a picture are typically relatively identical, the frame image is divided into blocks rather than storing each pixel separately, allowing the generally little variations between each pixel to be saved with fewer bits. Intra-frame prediction uses spatial redundancy or the correlation between pixels inside a single frame, to extrapolate prediction values from already coded pixels for efficient delta coding. Intra-coding is utilised in so-called intra-frames, which are temporally independent coded. Intra- and inter-frame prediction can be used in temporally coded anticipated frames (e.g. MPEG's P- and B-frames). The intra-class correlation coefficient is used to assess the PSNR and bit rate performance parameters. In general, prediction is lowered to 30.98% [6, 7].

In a video compression stream, an inter-frame is a frame that represents one or more neighbouring frames. The phrase's "inter" component alludes to inter-frame

prediction. This type of prediction attempts to improve compression rates by utilising temporal redundancy between neighbouring frames. A frame in a video compression stream that represents one or more neighbouring frames is known as an inter-frame. The "inter" component of the term refers to inter-frame prediction. This type of prediction tries to increase compression rates by leveraging temporal redundancy between adjacent frames. A block matching algorithm is used to finish this procedure. If the encoder's search is successful, the block can be encoded as a motion vector pointing to the spot.

Motion estimation is a motion vector estimate method that combines H.264/AVC bit rate reduction with RDO computation. Each input frame is divided into 16× 16 and 4 × 4 macro-blocks based on texture information (MB) [8].

The intra 16 × 16 intra-prediction mode predicts the whole 16 × 16 luma block and is perfect for coding smooth parts. When executing intra-prediction for a macro-block, H.264/AVC offers partition sizes ranging from 16 × 16 to 4 × 4. After a systematic study through reference frames, the prediction signal for each MxN luma block is defined using a translational motion vector and reference image index. To reduce the motion-compensated residue, the motion vector's resolution is increased to 1/4 pixel. H.264/AVC additionally supports intra-macro-block coding in addition to inter macro-block coding, resulting in higher PSNR values [9]. Each 4 × 4 block is predicted from geographically neighbouring data when utilising one of the nine intra 4 × 4 modes. This is ideal for encoding visual signals with a lot of information. Particle swarm optimization (PSO) has also been proposed to decrease the processing cost of hierarchical quad tree-based coding unit splitting. The PSO's chromosome and fitness function are successfully handled as the needed coding unit partitioning pattern for exhaustive partitioning and, respectively, the rate-distortion cost. In terms of PSNR, this produces the best results [10].

No reference is required to estimate the PSNR of digital video sequences encoded using the H.264/AVC technology. The goal is to replace a complete reference measure like PSNR (which needs both the original and processed video data) with the no metric that only works with the encoded bitstream. We may save a substantial number of computations necessary to decode the video pixel values because we are only working with the encoded bitstream [11].

More data are created as the number of smartphones increases, demanding compression, necessitating the creation of a new intra-prediction codec for the H.264 standard [12]. This method produces best results in terms of PSNR, compression ratio, and execution time. Our proposed method, known as best prediction matrix mode (BPMM), has overcome a number of problems. The huge overhead bits and the decoding method used to get the source video prior to projection are maybe the most critical difficulties. After overcoming these obstacles, (BPMM) underwent comprehensive testing and performance evaluations against the H.264 standard. Using pixel correlation analysis, the development of a CTU node model, a multi-core resource allocation method, and an intra/inter-frame joint WPP coding technique for multi-core platforms, researchers are able to increase compression whilst attaining good bit rate and PSNR results [13]. The performance research of the ITU-T H.264 standard,

including PSNR and MSE numbers to acquire a better knowledge of H.264 applications in industries and everyday life. Several test sequences are used to acquire a better understanding of the advantages of H.264 over previous video coding standards [14].

3 Proposed Work

Our implementation work shows the two different areas in intra- and inter-prediction compression of endoscopy videos. The block diagram for intra- and inter-prediction is depicted below (see Fig. 2.)

The endoscopy video is in the AVI format, which is a raw video direct sample used for compression and is based on 64 × 64 macro-blocks. For the evaluation of endoscopic video, popular ways for measuring the coding performance of a video coding standard include employing an empirical criterion such as peak signal-to-noise ratio (PSNR), compression ratio, and bit rate. Let's go over the most common definitions for these variables. As illustrated in Fig. 2, endoscopy video is made up of a series of video frames that are sent into the pre-processing step, where colour conversion takes place. The greyscale frames are then forwarded for block production of 16 × 16 to 64 × 64 bits after this phase. Then, the transformation and quantization take place, finally the encoder encode and generate compressed bits. The compressed bitstream is either saved or sent. As shown in Fig. 2, a video decoder decompresses the bitstream to produce a sequence of decoded frames. In our testing, 0–300 frames are created each 200–250 MB endoscopic video [15, 16]. In PSNR, p and q represent image sizes; $f(x, y)$ represents the original input pictures, and $g(x, y)$ represents the reconstructed image. The PSNR is expressed in decibels.

$$\mathrm{PSN} = 10\log_{10}\left[\frac{\frac{1}{pxq}\sum_{i=1}^{p}\sum_{j=1}^{q}(f(x,y))^2}{\frac{1}{pxq}\sum_{i=1}^{p}\sum_{j=1}^{q}(f(x,y)-g(x,y))^2}\right]$$

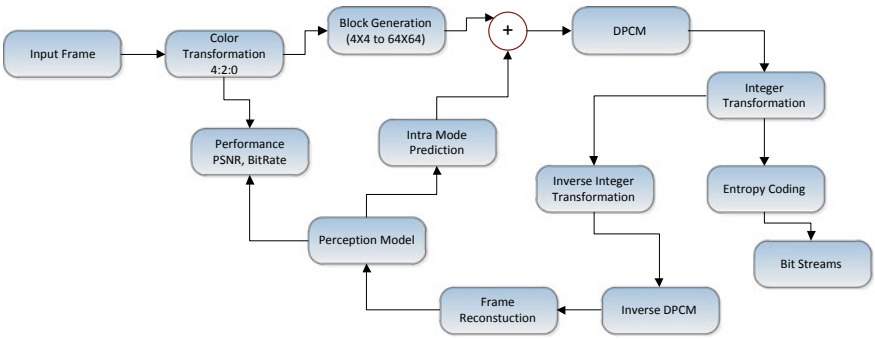


Fig. 2 Block diagram for intra- and inter-prediction for endoscopic video compression

Compression ratio defined as the ratio of the original image and compressed images.

$$\text{CR} = \frac{\text{Uncompressed}_{\text{image}}(\text{Bytes})}{\text{Compressed}_{\text{image}}(\text{Bytes})}$$

4 Experimental Results and Analysis

A parameter such as PSNR, bit rate, and compression ratio is taken into account for our study implementation. Because our research approaches are implemented in MATLAB, the experimental results demonstrated the efficacy of the suggested video compression methodology in terms of PSNR, bitrate, and compression ratio when compared to existing techniques. Let's have a look at both the intra-prediction and inter-prediction of endoscopic video implementations, which are available in our research finding (see Figs. 3 and 4) and shown as frames 1 and 2 for original and compressed video frames, respectively. We selected five samples of endoscopic films for this experiment, and the results are reported in Tables 1 and 2 for PSNR and bit rate in both the intra- and inter-prediction techniques. Our experiment yielded a compression ratio of 1:5 for both intra- and inter-prediction. PSNR is the method of determining the quality of a reconstructed picture. The PSNR value for a correctly compressed picture must always be low. The PSNR value for our suggested strategies in the execution of intra- and inter-prediction with respect to the values of findings in Tables 1 and 2 (see Figs. 5 and 6).

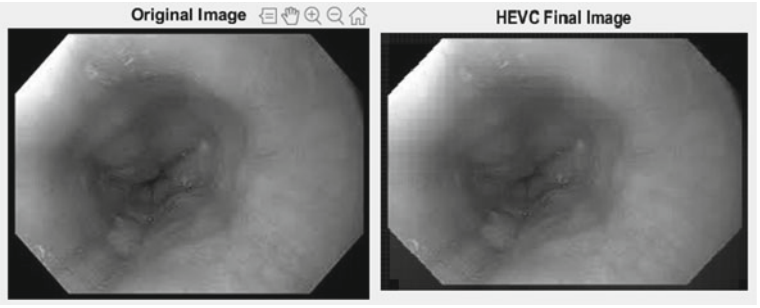


Fig. 3 Frame 1 of compression of image in intra-prediction

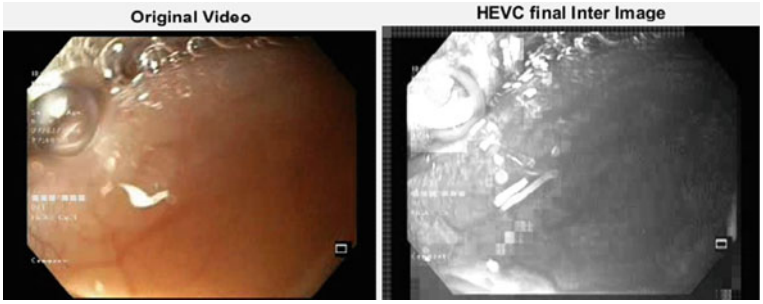


Fig. 4 Frame 2 of compression of image inter-prediction

Table 1 Result of intra-frame prediction

Endoscopic video	PSNR value (Db)	Bit rate in Mbps
1	39.58	0.051
2	41.47	0.076
3	44.68	0.065
4	45.35	0.077
5	50.01	0.078

Table 2 Results of inter-frame prediction

Endoscopic video	PSNR value (db)	Bit rate in Mbps
1	44.58	0.051
2	43.47	0.076
3	44.68	0.065
4	49.35	0.077
5	45.03	0.078

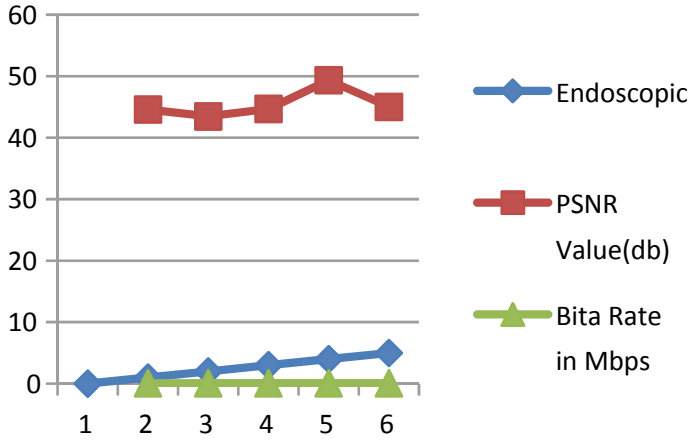


Fig. 5 PSNR and bit rate for intra-prediction

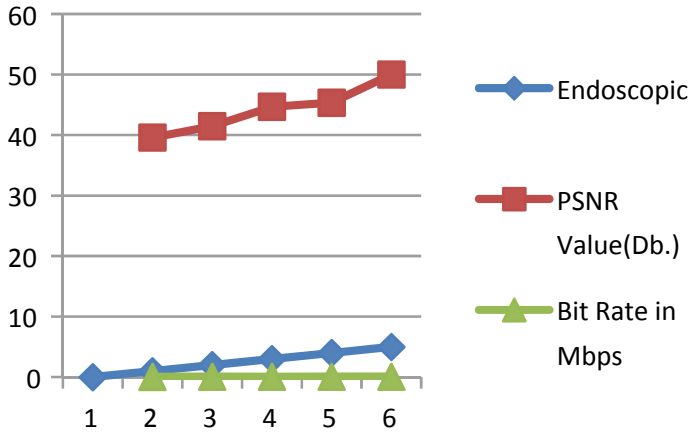


Fig. 6 PSNR and bit rate for inter-prediction

5 Conclusion

We compare the inter–intra-prediction system in this paper. The system has challenges such as the need to retain motion, maintain the quality of delivered images, and have a significant amount of storage capacity; these can be solved by the provided solutions. The image compression unit is a crucial component of an endoscopic system. We obtain an average PSNR improvement of 49–50 dB above the baseline H.264 profile for high-definition video with H.264 compatibility. Hence, with the aid of these proposed techniques, the problems in the existing image compression are reduced to a greater extent. Such prediction techniques can also be extended to the cloud-based application, which will be implemented in our future work.

References

1. Burnon D, Carly B, Ceelen W, De Roover A, Detry O, Duinslaeger M, Gys T, Hendriks J, Kolh Ph, Lamote J, Lerut J, Michel L (2011) The board of the royal belgian society of surgery by Cl. Bertrand. Endoscopy and surgery : a matter of diagnostic enlightenment and therapeutic liberty. Acta Chir Belg, 111:200–204
2. Gastrointestinal Endoscopy 48(3) (1998)
3. ISO 14971: 2007 Medical devices—application of risk management to medical devices. International organization for standardization. https://www.iso.org/standard/38193.html. Accessed 28 Sept 2019
4. Elakkiya S, Thivya KS (2021) Comprehensive review on lossy and lossless compression techniques. J Inst Eng India Ser B. https://doi.org/10.1007/s40031-021-00686-3
5. Jin X, Ngan KN, Zhu G Combined inter-intra prediction for high definition video coding
6. Shahid Z, Chaumont M, Puech W, Shahid Z, Chaumont M,· Puech W (2011) Signal image and video processing · January 2011 https://doi.org/10.1007/s11760-011-0225-9

7. Manjanaik N, Parameshachari BDb Hanumanthappa S Nc, Banud R (2017) Intra frame coding in advanced video coding standard (H.264) to obtain consistent PSNR and reduce bit rate for diagonal down left mode using gaussian pulse. IOP Conf Ser: Mater Sci Eng 225(2017):012209. https://doi.org/10.1088/1757-899X/225/1/012209
8. Mamatha RB, Keshaveni N (2018) Bit rate reduction For H.264/Avc video based on novel hexagon search algorithm. In: Mamatha RB, Keshaveni N (eds) Bit rate reduction for H.264/Avc video based on novel hexagon search algorithm. https://doi.org/10.21917/Ijivp.2018.0248
9. Park S (2012) PSNR-based Initial QP determination for low bit rate video coding. https://doi.org/10.6109/jicce.2012.10.3.315 print ISSN: 2234-8255 online ISSN: 2234-8883
10. Sakhare SV, Dalal UD (2020) Improving Psnr and processing speed for hevc using hybrid Pso for Intra frame prediction. Int J Multimedia Appl 12(1/2/3/4)
11. Slanina M, Rıcny V (2008) Estimating H.264/AVC video PSNR without reference using the artificial neural network approach. In: Proceedings of the international conference on signal processing and multimedia applications, pp 244–250. https://doi.org/10.5220/0001932502440250
12. Hamdy S, Zekry A, Mohamed WA (2019) Efficient H.264 intra frame CODEC with best prediction matrix mode algorithm. Int J Electron Telecommun 66(3):571–582 Manuscript received June 15, 2019; revised July, 2020. https://doi.org/10.24425/ijet.2020.134014.
13. Yin C, Hu D, Gu T (2018) Intra/inter-frame joint WPP coding algorithm on multicore platform. In: 3rd international conference on electrical, automation and mechanical engineering (EAME 2018) advances in engineering research, vol 127
14. Deshpande NR, Naik AK Performance analysis of advanced video coding (H. 264). Int J Eng Res Technol (IJERT) IJERT www.ijert.org ICONECT' 14 Conference Proceeding
15. Shenthil KKR, Nithyanandan L (2014) Medical video communication using modified HEVC over WiMAX network. In: Proceeding of 2014 international conference on communications and signal processing (ICCSP), IEEE, pp 808–812
16. Nandyal S, Gogi HK (2022) Implementation of compression technique for endoscopy video using intra coding HEVC. International conference on big data and cloud computing 2021-Springer

VREd: Virtual Reality in Engineering Education for Immersed and Interactive Learning: A Case Study

Anupama K. Ingale, J. Divya Udayan, and Savita P. Patil

Abstract Virtual reality (VR) is known for providing immersed user experience of 3D virtual world. In this paper, we have evaluated the learning of the engineering students and have also compared their learning with other the students who have not used VR. To conduct our study, we have developed a 3D model for various engineering fields such as, architecture, automobile and mechanical and so on. Further, we have also included VR interactions such as grabbing various 3D model, changing materials of the model and so on, to provide students an immersed experience of various 3D models of their respective discipline. Through the study, we have observed the students approximately 90% of students had good learning experience and have understood the concept better.

Keywords Engineering education · Virtual reality · Interactive learning

1 Introduction

Virtual reality (VR) and augmented reality (AR) have extensive applications in various fields such as, military, education, digital marketing, movies and so on. Moreover, the wide-range of head mounting devices in the market in affordable prices have made VR/AR even more popular. As world is facing the challenge due to novel COVID-19 in various sectors like health, education, trading and so on, with this, getting a quality education especially in higher education system has also a major concern. Most of the educational institutes have adopted online teaching platform

A. K. Ingale · S. P. Patil
Rajarambapu Institute of Technology, Isalampur, Maharashtra, India
e-mail: anupama.ingale@ritindia.edu
URL: https://www.ritindia.edu/

S. P. Patil
e-mail: savita.patil@ritindia.edu

J. Divya Udayan (✉)
Amrita Vishwa Vidyapeetham, Amritapuri Campus, Amritapuri, Kerala, India
e-mail: divyaudayanj@am.amrita.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_55

but the practical knowledge among the students is also very important especially in the field of automobile, mechanical and architecture engineering. The traditional education system needs to be reformed to cope up with the globalization era. To make graduates employable, it is necessary to transfer from traditional education system to learner centric approaches as in outcome-based education [1]. In addition to it, there are so many challenges such as funding, autonomy, one-size-fits-all approach and upcoming technology skill requirements in the future jobs designate teachers need to adopt new teaching methods for improving teaching learning process in education system. Incorporating technology into education allows to enhance learning methods and thus improves learning performance by creating and managing appropriate technological materials. Furthermore, this integration promotes student skills to learn how to use new technologies in their future life. Jobs and future requirements change at a fast speed, so students must be prepared to adapt to new environments and be proactive.

In this paper, we present a VR application for mechanical, civil and architecture engineering student. The application supports interactive learning, where student can interact with 3D objects and can also move around the 3D environment such as construction sites, various building plans and so on. We have also evaluated the effectiveness of the proposed system by conducting a comparison study among students. Our contribution in paper can be summarized as follows:

- Developed VR application for engineering students.
- Analysis of student learning in VR.
- Analysis of user experience in VR.

The paper is organized as follows; the next section explains related work followed by proposed methodology and then result and discussions.

1.1 Related Work

Recent advances in VR technologies have enabled low-cost head-mounted devices like Oculus Rift, HTC Vive [2], Microsoft HoloLens [3], VR headsets and so on to the market are in affordable price, thus making VR popular in various application fields like Military [4], education, engineering, medical, entertainment and so on. There are several research works carried out on VR for education and training [4, 5]. VR is used for architectural studies and have studied how student learning improved blending VR with architectural pedagogy [6]. To improving student's leadership and creativity, VR application is used in exploratory education [7]. VR applications are widely used in the medical field [8] so that students can have better experience and practice especially surgery. For dental education, improving students fine motor skills and hand-eye coordination is important for clinical setting. VR simulations are used to improve student's skills by providing them different practical sessions in VR simulator. It is also observed that the students learn fast and also practiced more

procedures and had achieved same level of competence as the traditional learning students [9, 10]. There are various 3D simulations for medical field such as DentSim [11], Simodont [12] and CadaVR [10], where students can practice procedure or lean the different parts of human bodies [13].

A part from medical field VR is extensively used in education [14–16], and the basic characteristic of VR immersion , interaction and imagination was presented by Pantelidis [17] and Freina and Ott [18] as 3 I's of VR. Abulrub et al. [19] have demonstrated how VR can be used in engineering field to apply theoretical knowledge to industrial problem thus making students industry ready. VR boosts the students confidence by allowing students to learn by doing [20] have presented how VR can benefit the engineering students [13] have developed a simple application for students to understand bubble sort algorithm and implemented a gamification technique so that student can learn concepts easily.

1.2 Methodology

Due to pandemic COVID-19, entire world adopted online teaching which impacted students learning especially in engineering field it is important that student has hands-on experience in their respective domain. It is challenging for teachers to simulate all the practical sessions particularly for civil, mechanical, automobile and architecture discipline. VR education can be viewed as new effective way of learning, where students not only experience immersed learning environment but also can interact with the 3D objects in the scene. Our proposed VR application supports multiplayer, where more than one player can enter into the application thus, allowing teacher and student to enter into the application at the same time. Therefore, teacher can monitor as well as explain student concepts leading to helping students to acquire good learning experience. Figure 1 shows the block diagram of our VR app.

VR scenes are created related to different domain, for example, for civil engineering discipline, we have created 5–6 VR scenes demonstrating different construction site, for architecture discipline, VR scenes of 3D buildings are created where students can walk through the building and have understanding of dimensions and different aspects related to architecture. These VR scenes are stored into the cloud and teacher has the access to these scenes where teacher can modify the scene as and when required. Further, all the interactions of students are tracked and teacher can view them, thus teacher can guide the student and observe their movements.

Application Module The main aim of the developed application is to provide a virtual classroom experience to the students, where students and teacher both can interact as in the real classroom. The advantage of the virtual classroom is the teacher can provide the visualization of concepts through various VR scenes and 3D object. We have developed various scenes using assets available in unity asset store, and Fig. 2 shows some of the scenes developed for the application.

The application consists of four main functional modules as shown in Fig. 3a:

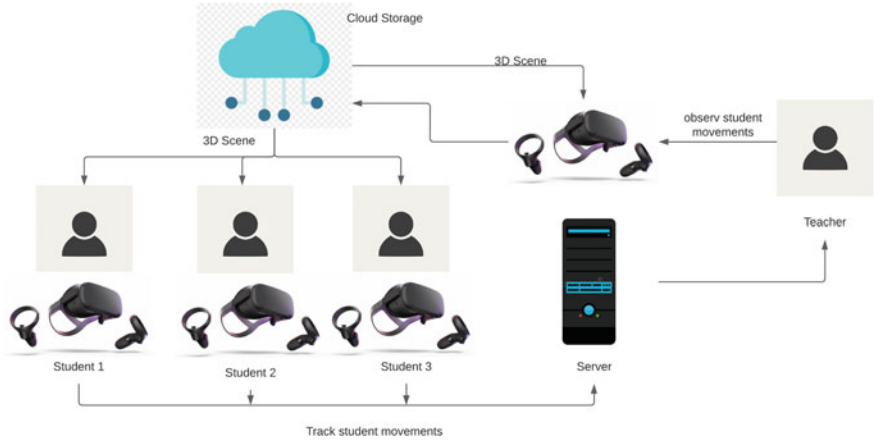


Fig. 1 System block diagram

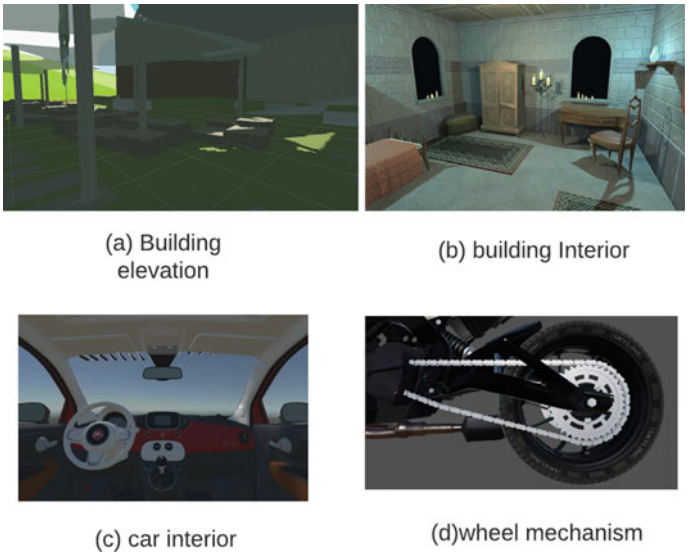


Fig. 2 Sample VR scenes

- User Manager: This module loads the teacher/ student module according to the login credentials. It also allows user to choose the virtual character and loads them in VR scene. This module gets the input from the login manager. Login manager module gives the VR scenes assess credential based on the type of user, for instance, teacher can have the assess to the VR scenes, where as student can view VR scenes.

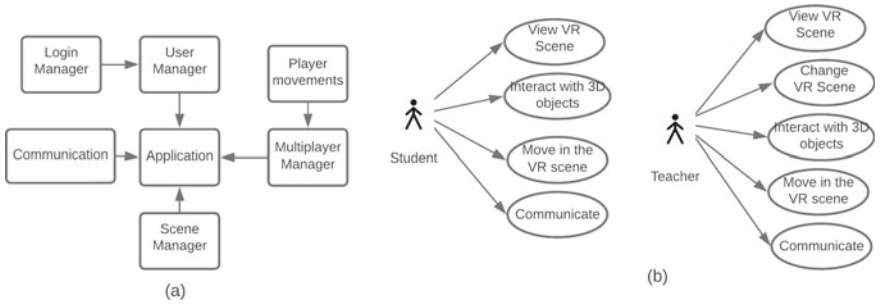


Fig. 3 **a** Application functional module **b** Use case diagram

- Communication: Voice on Internet module allows the user to communicate through the application, where user can talk with each other or listen to the video played by the teacher.
- Multiplayer manager: This module ensures the synchronization among multiple player, for instance, if one student picks up the object, another student can see in this app.
- Scene Manager: This module allows teacher to load various VR scenes stored in the cloud.

The application allows various interactions such as users can grab, throw or even they can change the properties of some objects. The interaction is implemented based on the needs of the students, for example, the architecture students can change the material of the wall, automobile students can grab the automobile parts to have better view of the automobile parts and so on. Two different functionalities are available based on the role of the users. If the user is teacher, then user can have accesses to various 3D scene and choose one which will be synchronized to all other user. They can also have access with video player within the VR scene, where the user can start or stop the video. The uses case diagram is shown in the Fig. 3b

2 Results and Discussions

The application is developed in Unity editor version 2020.3.12f1 with PUN2 package that support multiplayer, Oculus integration package to support interaction. The application was tested on group of 35 third year engineering students and 5 faculties, in which 2 faculties were from civil, 2 were from mechanical and 1 from automobile engineering discipline. Among 35 students, 15 students from civil discipline, 20 students from mechanical and automobile discipline. Before conducting the test, some general instructions were given for both faculties and students on how to use the application.

We conducted two survey one for analyzing user experience and the another for the learning experience. All questioner followed Likert scale: strongly agree, agree, somewhat agree, neutral, somewhat disagree, disagree and strongly disagree so that they are unbiased. After using the application, each of them was given feedback questioners as listed in Tables 1 and 2 which they were asked to fill after the test. To frame questioners, we have followed [21] further the user experience questionnaire is classified into three categories: interaction experience, immersed experience and multiplayer experience. The overall user experience was calculated as shown in Eqs. 1–4, where RQi denotes the response of the user for *i* th question from Table 1. To evaluate the effectiveness of the application, we have calculated average responses of interaction experiences, immersed experiences and multiplayer experience as shown in Eq. 4

$$R_{\text{interactionexperience}} = \frac{(\sum \text{RQ1} + \sum \text{RQ3} + \sum \text{RQ4}) - \sum \text{RQ3}}{4} \tag{1}$$

$$R_{\text{immersedexperience}} = \frac{(\sum \text{RQ5} + \sum \text{RQ6} + \sum \text{RQ6}) - (\sum \text{RQ8} + \sum \text{RQ9})}{5} \tag{2}$$

$$R_{\text{multiplayerexperience}} = \frac{(\sum \text{RQ10} + \sum \text{RQ11} + \sum \text{RQ12}) - (\sum \text{RQ13} + \sum \text{RQ14})}{5} \tag{3}$$

$$R_{\text{userexperience}} = \frac{R_{\text{interactionexperience}} + R_{\text{immersedexperience}} + R_{\text{multiplayerexperience}}}{3} \tag{4}$$

Figure 4a is the bar graph representing user responses collected to study user experience. From the graph, we can observe that 60% of the user found interaction easy, whereas 40% of user found interaction to be difficult. This is due to the users are unfamiliar with the VR usage and controller. Almost 90% of user had good immersed and multiplayer experiences. Figure 4b represents the bar graph of user responses collected for analyzing user leaning experience. From the chart, we can observe that average rating for Q1, Q3,Q4 and Q7 is approximately between the 4.5 and 5, whereas Q2 and Q6 are negative question thus lesser rating indicate that the user disagree for the negative questions, which indicates that the user enjoyed the learning in the VR.

3 Conclusion

Due to pandemic COVID-19, whole world faced challenges in various fields one among them was education sector particularly in the professional courses like engineering, medical and so on. In these courses, students should have hands-on experience in their domain. VR is used in various educational areas right from schools to medical sector. In this paper, we have presented our VR class room application and

Table 1 List of questionnaire for user experience

Q.No	Questionnaire
Interaction experience	
1	I was able to grab the object as I intended
2	I found it difficult to understand the how to use Virtual hands
3	I felt like I was using my own hands
4	My virtual hands are not coinciding with the 3D object
Immersed experience	
5	User menu was easy to operate
6	Overall look of virtual world was attractive
7	I was able to connect the subject concept with the VR scene
8	I felt too much distracted in virtual world
9	I found it difficult to interact with the other players
Multiplayer experience	
10	I was able to see the interaction of other players
11	I was able to listen to the instruction of the teacher clearly
12	I was able to exchange the object with other players
13	VR scene changes were not synchronized
14	Users voice was not clearly audible

Table 2 List of questionnaire for learning experience

Q. No	Questionnaire
1	I was able to relate the concepts well
2	I was more distracted and confused
3	Due to interaction, I was able to understand the working model
4	I was engaged well and was able to learn concepts easily
5	Guided instructions were helpful to understand the concept
6	I was not able to relate the topic with the virtual world scenes
7	VR has better learning as compared to video tutorials

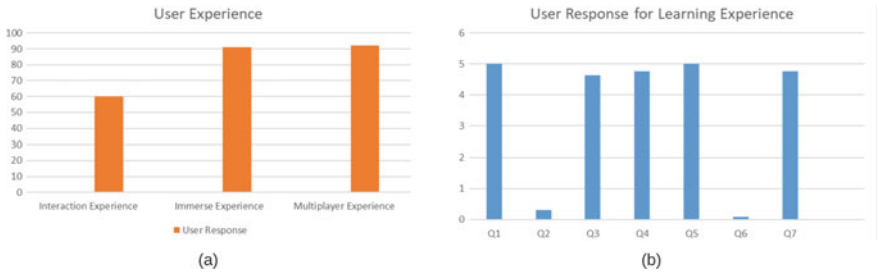


Fig. 4 **a** User experience **b** User learning experience

have conducted a case study on how students learning can be improved using high end technologies like VR. During our work, we faced difficulty in developing 3D scenes for specific concepts. Developing a 3D model for specific engineering domain like diesel engine, various parts of vehicles are still a time consuming task and 3D models are not freely available. Thus, creating the 3D study material especially for mechanical and automobile engineering will be our focus. In addition to this, we will also focus on adding features like virtual laboratory in which students can perform the laboratory experiments and thus get some hands-on experience in their domain.

Acknowledgements We express our gratitude to SEED GRANT, Rajarambapu Institute of Technology for enabling us to undertake this project, providing us with the required hardware and creating an environment that inculcates curiosity. We thank all the staff that have helped us in any way. We extend our thanks to our families for being supportive throughout this journey.

References

1. Gurukkal R (2020) Outcome-based education: an open framework. Higher Educ Fut 7(1):1–4
2. "HTC Vive." https://www.microsoft.com/en-us/hololen. Accessed: 03 Sept 2021
3. "Microsoft HoloLens." https://www.microsoft.com/en-us/hololen. Accessed: 03 Sept 2021
4. Balasundaram N, Ingale AK et al (2020) Virtual reality based therapy to mitigate ADHD. Psychol Educ J 57(9):1334–1341
5. Sala N (2021) Virtual reality, augmented reality, and mixed reality in education: a brief overview. Current and Prosp Appl Virtual Reality Higher Educ 48–73
6. Maghool SAH, Moeini SHI, Arefazar Y (2018) An educational application based on virtual reality technology for learning architectural details: challenges and benefits. ArchNet-IJAR: Int J Arch Res 12(3):246
7. Lin MT-Y, Wang J-S, Kuo H-M, Luo Y (2017) A study on the effect of virtual reality 3d exploratory education on students' creativity and leadership. Eurasia J Math Sci Technol Educ 13(7):3151–3161
8. Hamilton D, McKechnie J, Edgerton E, Wilson C (2021) Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. J Comput Educ 8(1):1–32
9. Roy E, Bakr MM, George R (2017) The need for virtual reality simulators in dental education: a review. Saudi Dental J 29(2):41–47
10. Makransky G, Lilleholt L (2018) A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educ Technol Res Devel 66(5):1141–1164
11. "DentSim." https://image-navigation.com/home-page/dentsim/. Accessed: 09 Apr 2021
12. "Simodont." https://www.simodontdentaltrainer.com/teachers-software/. Accessed: 03 Spet 2021
13. Nicola S, Virag I, Stoicu-Tivadar L (2017) Vr medical gamification for training and education. In: eHealth, pp 97–103
14. Chiou GF (1995) Learning rationales and virtual reality technology in education. J Educ Technol Syst 23(4):327–336
15. Jiang J, Zhi L, Xiong Z (2018) Application of virtual reality technology in education and teaching. In: 2018 International joint conference on information, media and engineering (ICIME), pp 300–302
16. Kavanagh S, Luxton-Reilly A, Wuensche B, Plimmer B (2017) A systematic review of virtual reality in education. Themes in Sci Technol Educ 10(2):85–119

17. Pantelidis VS (1997) Virtual reality and engineering education. Comput Appl Eng Educ 5(1):3–12
18. Freina L, Ott M (2015) A literature review on immersive virtual reality in education: state of the art and perspectives. Int Sci Conf Elearn Softw Educ 1:10–1007
19. Abulrub A-HG, Attridge AN, Williams MA (2011) Virtual reality in engineering education: the future of creative learning. In: 2011 IEEE global engineering education conference (EDUCON), pp 751–757
20. Stuchlíková L, Kósa A, Benko P, Juhász P (2017) Virtual reality versus reality in engineering education. In: 2017 15th international conference on emerging eLearning technologies and applications (ICETA), pp 1–6
21. Gonzalez-Franco M, Peck TC (2018) Avatar embodiment. towards a standardized questionnaire. Front Robot AI 5:74

Augmented Reality: Application of ICT Tools for Innovative Pedagogy

Suniti Dutt, Yash Singh, Aastha Singh, Akshay Kumar, Deva Harsha, and Dikshesh Kumar

Abstract With the everyday advancements in information communications technology (ICT), many technological developments have colossally affected education, and one such example is augmented reality (AR). The goal of this research paper is the visualization of the difficult concepts by utilizing the AR technology, which can be a means to innovative learning for the kids seeking elementary education. The paper highlights the importance of utilizing the idea of incorporating learning through AR, instead of letting the kids watch videos on screen. In the current paper, the innovative learning results have been shown picturizing a simple concept of the solar system working. Here, free 3D models of the planets are downloaded from Sketchfab, thereby converted into.glb form, imported, implemented, and developed in PlugXR software. Such applications of AR can envision creative minds to an ever-increasing extent of learning, and unlike the YouTube videos that provide the students with the knowledge, these AR experiences allow for the kids to remain active in their real surroundings.

Keywords Augmented reality (AR) · Education · Pedagogical tools · PlugXR · Virtual world

1 Introduction

Augmented reality (AR) innovation is the incorporation of computerized components which are imparted to this present reality continuously (genuine world information) and trace the natural circumstances which are existing in reality and applies to applications and platforms [1]. It is a unique experience of a verifiable environment in which real-world objects are augmented with PC-generated perceptual data, and this occurs across a variety of tactile modalities, including optical, aural, haptic, somatosensory, and olfactory. Today, in the instructing processes, numerous educators use data information communication technology (ICT) tools, innovation, and

S. Dutt (✉) · Y. Singh · A. Singh · A. Kumar · D. Harsha · D. Kumar
Chandigarh University, Punjab, India
e-mail: sunitidutt@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_56

media communications as one of the advancements in instruction; this is supposed to propel the nature of schooling in accordance with innovative advances [1]. In recent times, with the emanation of AR, the innovation has been getting analysts and instructors' consideration as a fascinating other option and invigorating approaches to making instructing and learning devices [2]. One immediate worldview in advanced education is the atonement of instructive assets toward a supposed "tech-savvy" crowd who, generally, have grown up encompassed by advanced innovation. The understudies entering advanced education today are requesting a high standard of instruction that integrates the advanced world in which they live. There is now and again a propensity to try to integrate futuristic innovations into all that is completed in a pioneering education organization yet it isn't advanced locals rather carefully mindful understudies that we ought to be creating [3]. To instruct understudies, an especially small kid, standing out, commitment, and inspiration are a difficult undertaking. Accordingly, with its one-of-a-kind highlights, AR innovation could be utilized as an alternative fascinating instructive conveyance strategy as it won't just prompt high inspiration in understudy yet in addition its instinctive and easy to use association strategy could be utilized as a consideration grabber learning instrument [2]. AR is an innovation that mixes computerized data with the data from physical-world conditions, empowering users to cooperate with virtual articles, and view the actual climate all the while [3].

2 Background of AR

The uses of AR technology in a variety of fields, including education, have recently received more attention. In the education sector, the early works of investigating capability of AR to help learning in light were crafted by Billinghurst, when he originally made an AR-based pop-up book known as the "MagicBook" [2]. In an instructive setting, AR offers some unmistakable benefits over standard showing rehearses as it can permit the conveyance of blended techniques where understudies have customary gaining from composed materials combined with perceptions of the cycles in question or give collaborations between genuine articles and the advanced representations [3]. AR is a perspective on virtual items and reality, simultaneously, where virtual data are added and coordinated into the actual world. AR goes from reality to computer generated reality, which is outright drenching [4]. In the midst of the rise of data gathering and analysis, one of the key purposes of augmented reality is to emphasize certain features of the real world, increase comprehension of those elements, and decide brilliant and open knowledge that can be applied to verifiable applications. Virtual reality (VR) and augmented reality permit vivid and visual encounters for the client. In VR, this will be by means of the utilization of a head-mounted show that is either an independent gadget or a gadget fastened to a PC that drives the perception equipment. In AR, the experience is not quite the same as VR as it permits the client to overlay computerized content in reality climate and collaborate with that substance. This offers some particular benefits over VR

as it sanctions the client experience to be shared among bunches as opposed to being a solitary client experience, and it likewise gives an open door to clients to blend learning styles when the AR is joined with more customary types of showing materials like texts or address slides. This blended methodology could give a strong instrument that fulfils numerous student styles, permits cooperative learning, and gives expanded extension to carry subjects to life in a manner that has not been imaginable previously [3].

3 Literature

3.1 AR in Education

There has been a recent surge in study in the use of AR technology in a variety of fields, including education. The education industry is one of the top three prospective areas of development for AR and VR technology, according to the XR Association's 2019 AR and VR Survey Report. Teachers may employ augmented reality tools to create fun and educational math curriculum, and interactive classes can make science more interesting [5]. Defines augmented reality (AR) as a method of combining the real and virtual worlds by using digital devices to add a virtual layer to real-world environments (Bower et al.). Such digital tools can be used by teachers as well as students to communicate, explicate, and learn while allowing their imaginations to run wild. Students may quickly detect the extra digital layer in their environment through the lens, but higher-level thinking occurs when they can apply their understanding of what they saw with the AR to the actual world after the AR has been withdrawn. Pochtoviuk et al. [4] show that presenting educational materials as augmented reality objects has a significant impact on children, altering facial expression development, attention, stimulating thinking, and increasing information understanding. The research looks at how augmented reality can be used in a variety of fields, including mathematics, anatomy, physics, chemistry, and architecture. For educational purposes, the Google Expeditions app claims to be able to make learning more engaging through unique tours where students can observe a range of objects, while the teacher talks about them. Rambli et al. [2] investigate the construction and evaluation of an augmented reality alphabet book, which is a book that teaches the alphabet to preschool children using augmented reality. Play, singing, dancing, theater, and the utilization of information and communication technology were all used to encourage participatory learning. The notion of augmented reality (AR), as well as its pedagogical foundations and impact on the learning process, is introduced in [6]. The authors produced "cleARmaths," a useful android app for teaching AR vector geometry. The benefits and satisfaction of the professors' and students' observations and outcomes were reviewed, proving AR's potential in math classes.

3.2 Pedagogical Approach on AR

Numerous works have previously highlighted the potential benefits of AR in education [2, 4, 5]. Earlier studies overlooked educational strategies, naïve to the notion that an intervention's effectiveness is influenced not only by the technical features of the technology but also by the pedagogical tactics utilized to implement it. Garzón et al. [7] adopt a pedagogical approach to AR, doing a quantitative meta-analysis of 46 empirical research to determine how pedagogical practices might affect AR's influence on the education sector in the learning environment. The impact of moderating factors on students' learning outcomes in AR therapies has also been investigated. According to the data, interventions that employed a collaborative pedagogical approach had the best impact. Milovanovic et al. [8] recommended modifying the application's functionality for educational use. The design studio is the most important part of an architect's training. The Luminous Planning Table, as well as the CAP VR and BenchWork gadgets, made design studio sessions go more smoothly. They have argued for and emphasized the need of employing a range of external representations modalities during studio sessions, as well as utilizing the potentials of virtual reality and augmented reality to boost architectural design learning.

3.3 Uses of AR in Other Fields

Prior to 2010, the bulk of AR applications were complicated, expensive, and had limited approachability and acceptance. In the recent years, incorporation of AR technologies into mobile devices has enhanced the amount of AR apps. Barrow et al. [3], Milovanovic et al. [8], and Soltani and Morice [9] discuss some of the more interesting and inventive uses of augmented reality technologies that are already bringing benefits to a range of businesses. The goal of [9] was to define and comprehend the advantages of augmented reality in sports training and education. Using the Scopus, Pubmed, Web of Science, and SportDiscus databases, the findings were evaluated for their relevance to sports (practitioner, spectator, and customer). They have employed a variety of sports, such as basketball, rock climbing, football, and dodgeball, to help practitioners and spectators by teaching sports skills, offering additional information and feedback, stimulating practice, interacting with sporting events, and expanding customers' options. A variety of augmented reality learning and feedback systems have been discovered to be useful. Milovanovic et al. [8] provide an overview of virtual reality and augmented reality technologies based on a corpus of materials from literature available on computer-aided architectural design (CAAD). A careful examination of a few particular research projects reveals their strengths and faults, establishing benchmarks for future challenges. In the architecture domain, AR applications can be created using a variety of techniques, including HMD AR, Tangible AR, smart device AR (SDAR), and SAR, all of which involve the blending of the

actual environment with virtual data through various methods. The CAP VR environment has been shown to help students improve their design quality. Immersion in a virtual model of their design enhanced spatial cognition and improved design evaluation. Barrow et al. [3] present the results of a survey on the use of digital visualization technologies in the teaching of life science disciplines, as well as the development and testing of an augmented reality application to aid in the teaching of metabolism, namely glucose and insulin signaling.

4 Problem Definition and Proposed Design/Work

The students go up against an assortment of difficulties in their learning cycle every day, including trouble grasping theoretical galactic information, as well as an absence of creative vision and innovativeness in understanding cosmic peculiarities. In customary instructing, a class of understudies are mostly educated without respect for their ground knowledge of the concept. In the world, the solar system is an enormous and interesting framework, and the teachers generally end up giving just a couple of moments to the discussion. Students find it strenuous to apprehend on the grounds, concepts that they can't envision or adhere to in their day-to-day routines. Thus, they require a platform to help them in creating more prominent, inventive capacities. Augmented reality (AR) has its own true potential in the education field since it can give a consistent collaboration among real and virtual items. Subsequently, in this venture, we endeavor to create a platform that joins an application for a galactic learning with increased reality procedure. The literary, graphical, sound, and visual data superimpose on the understudy's ongoing environment. Such application empowers the students to effectively partake and collaborate with computer created recreations. We present an application that contains exhilarating ways of getting kids engaged with finding out about the astronomical world starting with the planets of the solar system.

The proposed product emphasizes effective learning methods. This research intends to develop a prototype that will assist teachers in teaching solar system ideas to children who are unable to grasp them. A similar R&D approach is used in this prototype where no coding has been used. Analysis, design, development, implementation, and assessment are all part of this technique (evaluation). Figure 1 shows the design methodology.

For developing AR-based solar system content as a learning platform for learning about planets, this has been done in the following stages: Analysis where the Google ARcore Web site was referred to find out performance-based analysis, and information for need-based analysis was gathered by using Google Forms to find out what elementary school kids needed. At the time of Design, in PlugXR, a folder for the surface tracking program was made. In the application, two different platforms were created: one for planets and one for the solar system. A variety of planet templates were created. At the stage of development, the planet's 3D models were downloaded via the Sketchfab app, which includes AR-related 3D models. Then, using Internet

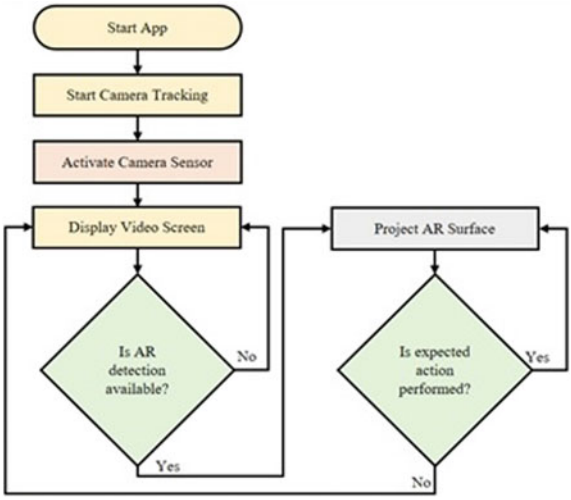


Fig. 1 Simple algorithm of the proposed AR system

converters, it was converted to.glb format. These 3D models were then imported into PlugXR's surface tracker, where the software's inherent textures were applied to the 3D models. As a result, the scale, rotation, and light direction were adjusted subsequently. It was spiced up with different texts and font colors to make it appealing to a primary school student. The content for templates was gathered, which were subsequently bound with planets and loaded into the platform. The animation scenes and touch actions were then added to these 3D models to provide some more functionality. Finally, while implementation, the final application was deployed by saving the content on the platform and publishing the final product. Then, during the setup stage of the application, a decent user interface was given through the PlugXR software, and hence, a similar android application was created. Evaluation was hence done by generating a link and a quick response (QR) code to download the final product.

5 Results

The result of this study is a learning media platform for the solar system for primary school kids. The simulation part of the software has been discussed further, using the images of the application made. Firstly, the application is launched on an android-based smartphone, and then, the menu screen appears (Fig. 2a) showing 2 options to select, "Solar System" and "Planets," either of which could be selected.

As this AR model requires a flat surface, on selecting any one of the option from (Fig. 2a), you would then be asked to move your device slowly (Fig. 2b), where you would be required to move the device slowly so that it can detect a flat surface. On detecting the flat surface, in case you have selected the option of "Solar system" icon, the particular AR model appears as seen in (Fig. 3). The animation (revolution

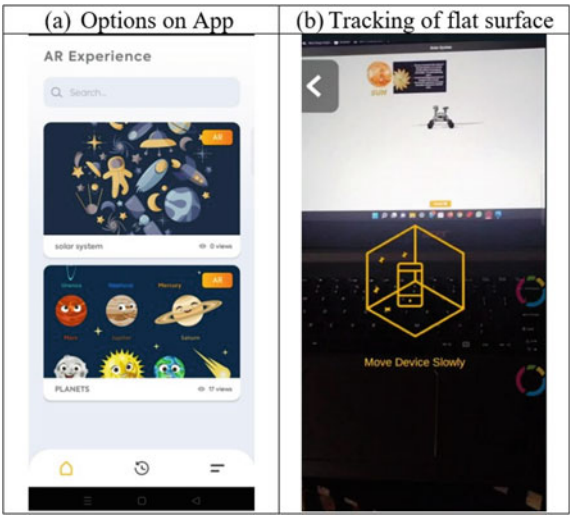


Fig. 2 Standard interface of the app

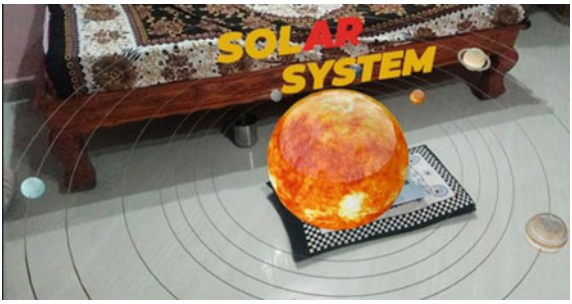


Fig. 3 AR experience of solar system (backgrounds are real environment)

and rotation) of the planet would hence starts with a single tap on the screen of your smartphone. If you select the option of "Planets," it shall display for you the information of the planets (Fig. 4a, b). These slides of information would part general information about the planets as per the requirements of a primary school kid. This user-friendly application can be used by kids themselves, provided they are familiar with the basic working of a smartphone.

6 Conclusion and Future Scope

In view of the consequences of the exploration and analyzation that have been directed, it tends to be presumed that AR is an innovation that is valuable in the educating and learning process. With the utilization of AR technology, it is anticipated to help the understudies as well as the educators to interface all the more

Fig. 4 AR experience of the different planets with their information (backgrounds are real environment)

genuine and intuitive with virtual articles. In the concluded application, the planets revolving around the sun appear to be a 3D environment (Fig. 3), whereas the second interface where the brief information of planets (Fig. 4) is given appears to be a 2D environment.

There remains to be some limitations which is discussed below. (1) Since this software is not an offline application, a fast Internet connection is essential. (2) The application, having built on the basic free version platform of PlugXR, works only on android operating systems. (3) The application built is a coding-free software; hence, features like asteroids and its revolution speed could not be added. Regardless of these flaws, PlugXR appears to be stable and should be user-friendly for primary school children. The application is a prototype that holds the capability of adapting new and advanced features that can be modified using various other codes and concepts, according to the requirements and the feedback from the end-users. The organization of the application on android platform moreover implies the uptake of the application has a wide reach. To further advance, the work will focus on iOS application development.

References

1. Tresnawati D, Fatimah DDS, Rayahu S (2019) The introduction of solar system using augmented reality technology. J Phys: Conf Ser 1402(7):1–7

2. Rambli D, Matcha W, Sulaiman S (2013) Fun learning with AR alphabet book for preschool children. Proc Comput Sci 25:211–219
3. Barrow J, Forker C, Sands A, O'Hare D, Hurst W (2019) Augmented reality for enhancing life science education. In: The fourth international conference on applications and systems of visual paradigms, Italy
4. Pochtoviuk SI, Vakaliuk TA, Pikilnyak AV (2020) Possibilities of application of augmented reality in different branches of education. Educ Dimens 54(2):92–106
5. Leighton LJ, Crompton H (2017) Augmented reality in K-12 education. Mobile technologies and augmented reality in open education. Chapter 14. 281–290
6. Schutera S, Schnierle M, Wu M, Pertzel T, Seybold J, Bauer P, Teutscher D, Raedle M, Heß-Mohr N, Röck S, Krause MJ (2021) On the potential of augmented reality for mathematics teaching with the application cleARmaths. Educ Sci 11:368
7. Garzón J, Kinshuk, Baldiris S, Gutiérrez J, Pavón J (2020) How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. Educ Res Rev 31
8. Milovanovic J, Moreau G, Siret D, Miguet F (2017) Virtual and augmented reality in architectural design and education: an immersive multimodal platform to support architectural pedagogy. In: 17th international conference, CAAD futures
9. Soltani P, Morice AHP (2020) Augmented reality tools for sports education and training. Comput Educ

Enhanced Gradual-N-Justification Methodology with Local Outlier Factor (LOF) for Hardware Trojan Detection

Pratyusha Robbi, M. Priyatharishini, and M. Nirmala Devi

Abstract Security and trust of any electronic system is one of the major concern in the present era of globalization. The malicious or unexpected modifications capable enough of accessing the system to change its functionality and weaken the system: Hardware Trojans. However, a methodology that is reference-free, proficient, and experience high-false positives is greatly encouraged. The study involves: a reference-free hardware detection analysis, gradual-N-justification methodology, and local outlier factor. GNJ methodology is an extensible linear algorithm that produces a list of suspicious signals, detects the HT, and reduces the false-positive rates. Local outlier factor adds a depth to the study by classifying the data points based on the local density, reduces the suspicious signals, and subsequently the number of iterations required to bring out the most suspicious signals. The proposed methodology does not fail to bring out any non-maskable Trojan if inserted into the circuit. Therefore, the proposed methodology extracts the most suspicious signals from a list of suspicious signals with high accuracy and less time.

Keywords Hardware Trojans (HT) · Hardware security · Malicious modifications · Gradual-N-justification (GNJ) · Suspicious signals (SS) · Local outlier factor (LOF) · Controllability · Observability · Machine learning · Unsupervised clustering

P. Robbi (✉) · M. Priyatharishini · M. Nirmala Devi
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: cb.en.p2vld20021@cb.students.amrita.edu

M. Priyatharishini
e-mail: mpriyatharishini@cb.amrita.edu

M. Nirmala Devi
e-mail: mnirmala@cb.amrita.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_57

1 Introduction

In the present era of globalization, more complex integrated circuit (IC) is coming into existence; companies are finding it difficult to fabricate designs with small feature size. This results in redistribution of IC design and functioning parts. Redistribution of certain parts can enable the malicious modifications to take place. It can be observed that these the entire chain of IC is being subjected to the malicious manipulations [1] without the knowledge of the company. There are a number of malicious modifications or manipulations that can get unauthorized access to the system, hence can weaken the system. Weakening can be either degrading the systems performance or trustability [2] issues or effects the data rectitude. Alteration in systems functionality due to the insertion, deletion, or modification of maligns is known as hardware Trojans. Hardware Trojans is a threat model that can pose a hitch when triggered or activated in any system in general.

HT can be inserted at both preprocessing and post-processing stage of silicon. When system performance records abnormalities, detecting or testing for presence of HT can be done either in run-time or design-time (preprocessing stage) and in test-time for post-processing stage. Reverse engineering (RE) [3], side-channel analysis (SCA) are some of the most common methods used to gain access to the data, and hence, security of the system is lost. The main objective of any manufacturer is to deny the access acquired by the HT using reverse engineering or side-channel analysis techniques. Manufacturer or the company have to make sure that system performance is not degraded and have to find alternative approaches to detect the HT or to determine a secure system that remains ineffective even after the insertion of HT [4].

Reverse-engineering technique is a destructive methodology which would consume time when a higher degree of circuits is required. On the similar grounds, detection of hardware Trojans using side-channel analysis [5] becomes exhaustive process due to process variations, temperature variations, and if a reference circuit becomes necessary in the process of circuit functionality analysis. Therefore, it can be observed that the surviving methods to detect HT prove to be inefficient due to its complexity and non-climbable property [6].

This work provides a depth analysis to the existing GNJ technique to detect hardware Trojans. Extending the GNJ technique by incorporating the local density-based outlier detection methodology [7] enhances the detection process and promises to bring out true suspicious signals at the end.

Further sections are organized as follows: In Sect. 2, an overview of background study is given. In Sect. 3, motivation and contribution are discussed to enhance the already existing methodology. In Sect. 4, a detailed illustration of the proposed methodology is highlighted. In Sect. 5, experimental results are analyzed, and conclusion of the work is discussed in Sect. 6.

2 Background Study

The abovementioned limitations [8] in the existing methodologies can be suppressed by embracing machine learning techniques where the necessity of golden or reference circuit can be avoided. ML techniques are classified as: supervised [9] and unsupervised machine learning techniques. Supervised machine learning technique involves: linear regression, polynomial regression, regression trees, nonlinear regression. Some of unsupervised machine learning methods are as follows: K-nearest neighbor, K-means [10], neural networks, anomaly detection, hierarchical clustering, principal component analysis, and independent component analysis.

Supervised machine learning [11] techniques require both input trained data along with the corresponding mapped output data. Unsupervised machine learning is a technique which does not require any training datasets to model the data. It is capable enough to find the hidden parts and insights from the dataset. The input data of the unsupervised algorithm cannot be directly applied to regression or classification because the data provided as input have no corresponding mapped output data. This technique main objective is to determine the underlying structure of the dataset, classify based on the properties, and expose it in a compressed format yet an easy to look at format.

The gradual-N-justification methodology [12] uses K-means clustering algorithm. This can be applied to the output produced by a reference-free HT detection methodology. GNJ processes the list of signals reported as suspicious and produces a list of signals that are considered as most suspicious [13]. This is an iterative process where the justification effort is determined by the parameter N. Adoption of unsupervised ML enables the freedom to detect anomalies, and segmentation in K-means can be done effortlessly. GNJ technique can be computed in a straightforward fashion as the clustering process involved eliminates the complexity of labeling the input data and has the capability to unfold the hidden key segments.

3 Motivation and Contributions

The proposed GNJ technique has a higher time complexity as it is a iterative process and would terminate only when the list of signals obtained as output are outliers or have similar densities. Introducing a slight tweak to GNJ, can reduce the time required for computation. Outlier detection algorithm can be proven as advantageous because it can help treating inliers and outliers, can further process only the outliers (major concern). Local outlier factor is an outlier detection technique that adds a depth analysis to the already existing gradual-N-justification methodology.

The list of output signals at the GNJ technique can be exercised to obtain a list of most suspicious signals by utilizing local density. Final classification is done based on the anomaly score either as outlier or inlier. The entire developed framework is

an iterative procedure (climbable procedure) but has lesser time complexity when compared to the GNJ.

The proposed methodology underlines the following features:

1. A methodology that combines signal justification, gradual-N-justification technique, and outlier detection technique,
2. Introducing a local outlier factor, an outlier detection methodology that can be applied to a list of SSs produced by the GNJ technique,
3. Inertia and LOF score (anomaly score) as constraints to check and verify at various stages of the proposed iterative process,
4. Performing evaluation on the trusted ISCAS'85 combinational circuits, and
5. Demonstrating a relationship with respect to GNJ methodology and local outlier factor.

4 Methodology

The proposed methodology can be divided into three sub-algorithms: (a) Controllability and observability for Trojan detection (COTD), (b) Gradual-N-justification methodology [signal justification and unsupervised machine learning algorithm], and (c) Outlier detection methodology.

4.1 COTD Analysis

Algorithm 1: Controllability and observability for Trojan detection, a reference-free HT detection method. Classification as HT-free or HT-inserted is made based on SCOAP [14] controllability and observability values. A net with high CC, calculated using (Eq. 1),

$$CC = \left(CC0^2 + CC1^2\right)^1/2 \tag{1}$$

or high CO [15] is treated as nets where controlling and observing is very difficult; hence, these are the ones that need further analysis and are displayed as a list of reported suspicious signals (Refer Fig. 1a).

4.2 Gradual-N-Justification Methodology

Algorithm 2: Gradual-N-justification (GNJ) methodology uses two fundamental concepts: signal justification and unsupervised ML clustering algorithm: K-means. Determining the number of realizable input combinations to set the output signal or net to 1 or 0 is called as signal justification process. K-means clustering is one

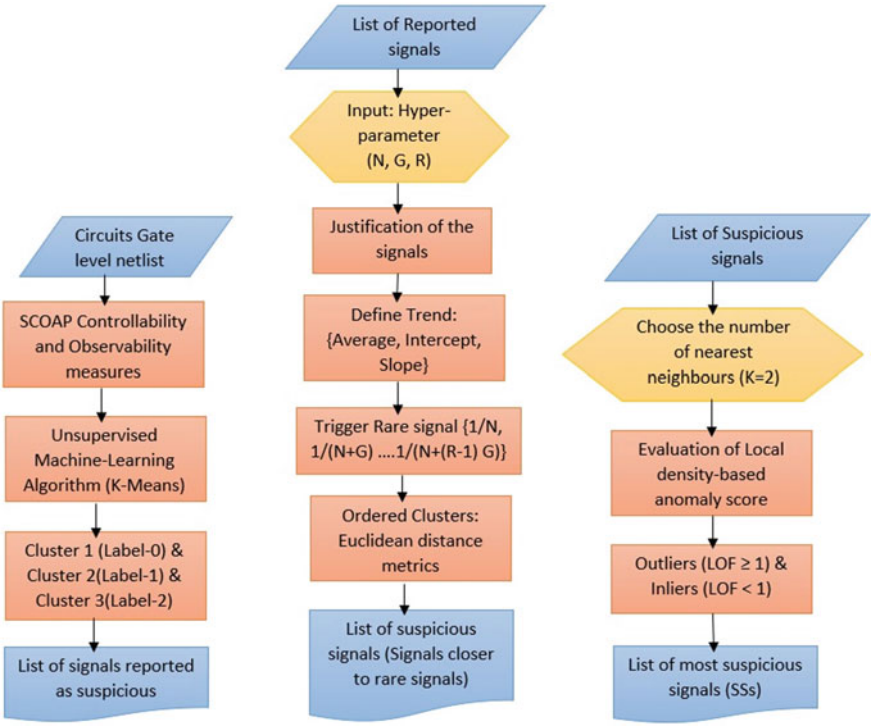


Fig. 1 Implementation of COTD, GNJ, and LOF methodologies

of the standard clustering algorithms where centroid-based clustering is used. GNJ methodology uses three parameters (N, G, R) to ascertain the incremental gradual justification process. N defines the effort required for justification process. G and R control the justification process growth. The break condition is determined by the inertia, a performance metrics of the K-means clustering (Refer Fig. 1b).

4.3 Outlier Detection Methodology

Algorithm 3: An unsupervised ML technique where distinguishing a data point or signal is done based on local density deviations: anomaly score. Basic principle used in this density-based approach is KNN algorithm. The computed anomaly score is termed as local outlier factor and is calculated for every signal. Signal having density similar to its neighbors: inliers (within the same cluster). Signal having densities lesser or different densities with respect to its neighbors, outliers (Fig. 1c).

Table 1 Illustration of ISCAS'85 circuits

Circuit	# Gates	#Inputs	#Outputs	#rSS	#mSS	#iterations(LOF)	#iter
C17	4	5	2	9	4	2	2
C432	153	36	7	133	4	2	3
C880	357	60	26	560	8	2	4
C1355	514	41	32	1195	4	3	5
C1908	855	33	25	1421	4	3	5
C3540	1179	50	22	493	8	2	4
C5315	1726	178	123	1661	5	3	5
C6288	2384	32	32	3683	5	3	6

5 Experimental Results

Detecting HT on such a large number of inputs, outputs, and intermediate nets can be more exhaustive. Applying the COTD methodology to spot-light signals based on (high priority) CO and CC values reduces the effort and time complexity, brings out accurate results. Figure 3 demonstrates different clusters obtained due to the adoption of COTD technique. Table 1 portraits the specifications of the circuits considered, ISCAS'85 [16].

To attain the similar list of most suspicious signals, the justification effort is less for the proposed work when compared to that of the already existing method. For example, consider the analysis for C1355 (Fig. 4a) performed using the proposed work algorithm (Fig. 2); justification effort (N) ranges from (15, 21, 27) with a periodic increment of 6 (i.e., $G = 2$ and $R = 3$), where G and R are known to limit the growth of justification effort. Whereas, N ranges from (15, 21, 27, 33, 39) in the case of the already existing methodology. Similarly, consider the analysis made for C6288 (Fig. 4d); N ranges from (15, 21, 27) with a periodic increment of 6 (i.e., $G = 2$ and $R = 3$), where G and R are known to limit the growth of justification effort. Whereas, N ranges from (15, 21, 27, 33, 39, 45) in the case of the already existing methodology.

Some of the circuit's analysis have been displayed in Fig. 4. From the analysis, it is clear that though the proposed work has repeatable implementation, the circuit is undergoing a smaller number of iterations (defined by the justification effort, Fig. 2). Limited number of iterations has a direct impact on the time complexity factor. Hence, an improvement in the time complexity is noticed.

6 Conclusion

The performance of local outlier factor, an outlier detection scheme is spotlighted in Fig. 5. The number of iterations taken to distinguish the most suspicious signals from a

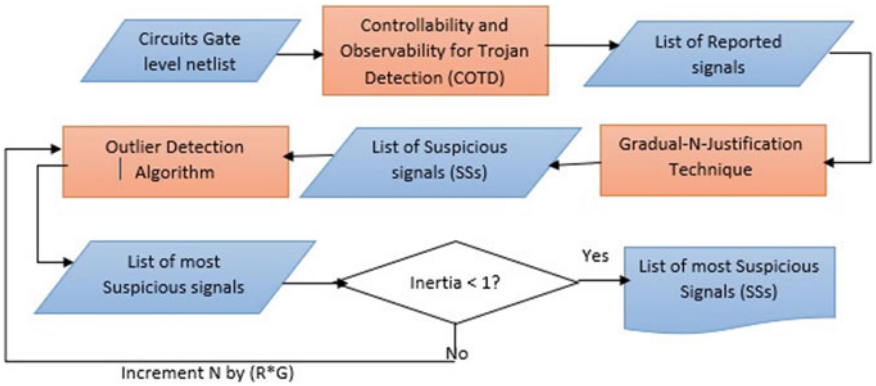


Fig. 2 Proposed methodology

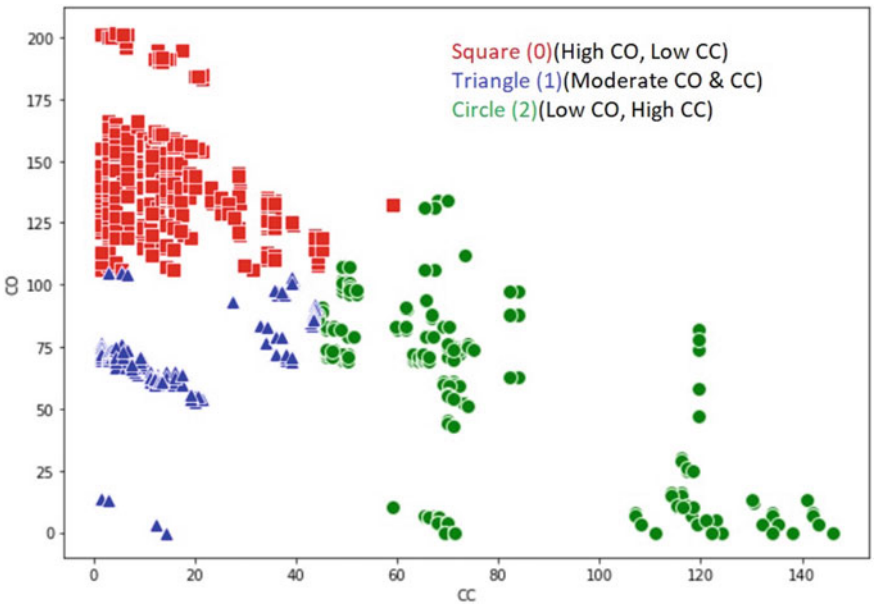


Fig. 3 COTD: CC versus CO

list of signals reported as suspicious based on controllability and observability values (initial) is displayed in Fig. 5. As the circuit complexity increases, it is evident that with less iterations, the desired list of most suspicious signals can be obtained (Table 1). A significant decrease in the iterations can subsequently result in diminished time complexity even when higher ordered circuits are taken into consideration. With a reduction in required effort, high rates of true positive rates are observed due to the insertion of an unsupervised ML-based outlier detection technique. Any inserted

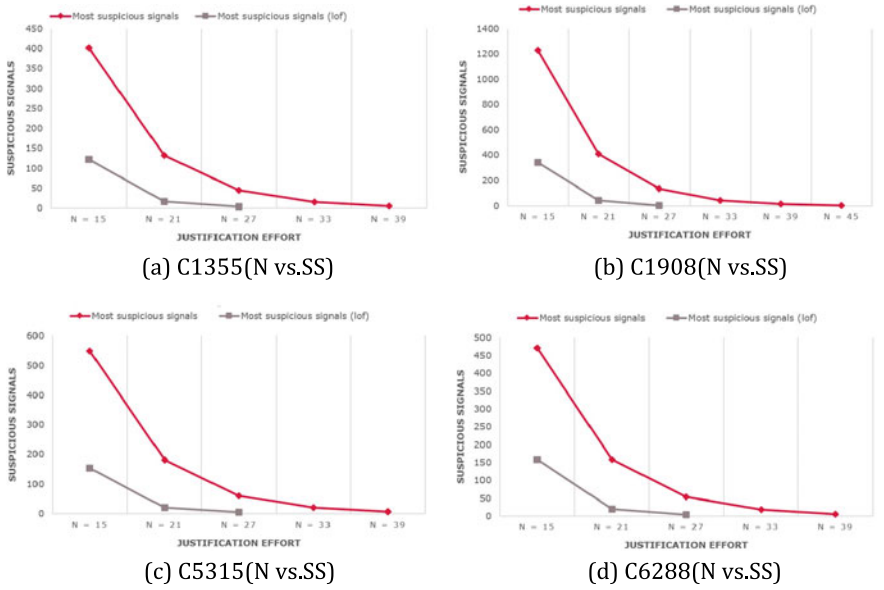


Fig. 4 Analysis of ISCAS'85 circuits

or detected HT are marked as most suspicious signals at the end of this iterative process. Escalated speed and improved systems performance can be perceived with the proposed work.

Some of the advantages of the proposed methodology: A reduction in the execution time is observed. The requirement of golden circuit or Trojan-free circuit is eliminated to detect the Trojans, if exists. The use of unsupervised machine learning technique

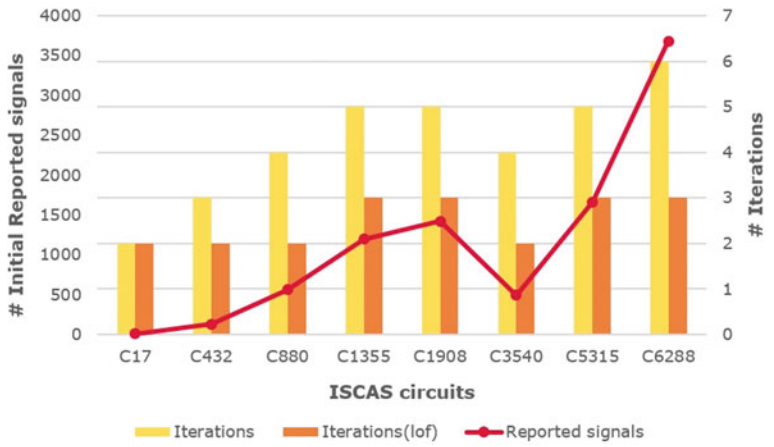


Fig. 5 Improvements of using proposed methodology

eradicates the necessity of labeling the data and makes the clustering process more effective, which plays a crucial role classification of Trojan-free and Trojan-inserted nets. The proposed methodology limits the detection of HTs that are masked by the circuit's internal functionality.

As future scope of work, the proposed methodology can be extended to sequential circuits. The proposed methodology is evaluated only in the presence of combinational HTs. At later stages, the HT's insertion can also expand till sequential HTs.

References

1. Priyatharishini M, Nirmala Devi M (2018) Detection of malicious circuit in hardware using compressive sensing algorithm. In: 2018 Second international conference on advances in electronics, computers and communications (ICAECC), pp 1–5. https://doi.org/10.1109/ICAECC.2018.8479492
2. Salmani H (2018) Trusted digital circuits—hardware Trojan vulnerabilities, preventionand detection. Springer, Cham, Switzerland
3. Bao C, Forte D, Srivastava A (2016) On reverse engineering-based hardware Trojan detection. IEEE Trans Comput Aided Des Integr Circuits Syst 35(1):49–57. https://doi.org/10.1109/TCAD.2015.2488495
4. Saran T, Ranjani RS, Devi MN (2017) A region based fingerprinting forhardware Trojan detection and diagnosis. In: 2017 4th international conference on signal processing and integrated networks (SPIN), pp 166–172. https://doi.org/10.1109/SPIN.2017.8049937
5. Hu T, Wu L, Zhang X, Yin Y, Yang Y (2019) Hardware Trojan detection combine with machine learning: an SVM-based detection approach. In: 2019 IEEE 13th international conference on anti-counterfeiting, security, and identification (ASID), pp 202–206. https://doi.org/10.1109/ICASID.2019.8924992
6. Krishna AV, Prabhu E (2022) Hardware Trojan detection at behavioral level using inline assertions and verification using UVM. In: Smys S, Balas VE, Palanisamy R (eds) Inventive computation and information technologies. Lecture notes in networks and systems, vol 336. Springer, Singapore
7. Budiarto EH, Erna Permanasari A, Fauziati S (2019) Unsupervised anomaly detection using K-means, local outlier factor and one class SVM. In: 2019 5th international conference on science and technology (ICST), pp 1–5. https://doi.org/10.1109/ICST47872.2019.9166366
8. Bhunia S, Hsiao MS, Banga M, Narasimhan S (2014) Hardware Trojan attacks: threat analysis and countermeasures. Proc IEEE 102(8):1229–1247. https://doi.org/10.1109/JPROC.2014.2334493
9. Harsha GMKS, Nikhil J, Eswar MS, Ramesh SR (2021) Hardware Trojan detection using supervised machine learning. In: 2021 6th international conference on communication and electronics systems (ICCES), pp 1451–1456. https://doi.org/10.1109/ICCES51350.2021.9489081
10. Wang Z, Zhou Y, Li G (2020) Anomaly detection by using streaming K-meansand batch K-means. In: 2020 5th IEEE international conference on big data analytics (ICBDA), pp 11–17. https://doi.org/10.1109/ICBDA49040.2020.9101212
11. Kulkarni YP, Mohsenin T (2016) Adaptive real-time Trojan detectionframework through machine learning. In: 2016 IEEE international symposium on hardware oriented security and trust (HOST), pp 120–123. https://doi.org/10.1109/HST.2016.7495568
12. Salmani H (2022) Gradual-N-Justification (GNJ) to reduce false-positive hardware Trojan detection in gate-level Netlist. IEEE Trans Very Large Scale Int (VLSI) Systems 30(4):515–525. https://doi.org/10.1109/TVLSI.2022.3143349

13. Jianxin W, Boren W, Ming Q, Lei Z, Chaoen X (2016) Hardware Trojan detectionbased on the distance discrimination. In: 2016 First IEEE international conference on computer communication and the internet (ICCCI), pp 404–407.https://doi.org/10.1109/CCI.2016.7778953
14. Kok CH, Ooi CY, Moghbel M, Ismail N, Choo HS, Inoue M (2019) Classification of Trojan nets based on SCOAP values using supervised learning. In: 2019 IEEE international symposium on circuits and systems (ISCAS), pp 1–5.https://doi.org/10.1109/ISCAS.2019.8702462
15. Salmani H (2017) COTD: reference-free hardware Trojan detection and recoverybased on controllability and observability in gate-level netlist. IEEE Trans Inf Forensics Secur 12(2):338–350. https://doi.org/10.1109/TIFS.2016.2613842
16. ISCAS. Accessed: April. 23, 2022. [Online]. Available: https://pld.ttu.ee/~maksim/benchmarks/

IoT-Based Automatic Irrigation Scheduling Using MQTT Protocol

Arunava Laha, Bajradeepon Saha, Aishwarya Banerjee, Pratap Karmakar, Debaprasad Mukherjee, and Arpita Mukherjee

Abstract In this paper, an IoT-based wireless sensor network for automatic irrigation of agricultural fields has been designed and developed. Four numbers of sensor nodes are placed over the desired locations to collect data from an agricultural field, and a controller node is placed near the pump to control the water flow of the agricultural field. The sensor node consists of a microcontroller NodeMCU ESP8266 and a soil moisture sensor, and the controller node is made of NodeMCU ESP8266 and a relay to control the on–off state of the pump. The soil moisture sensor of these nodes collects soil moisture content data and delivers it to the cloud or server using the message queuing telemetry transport (MQTT) protocol, through wireless communications modules. The ThingSpeak cloud platform is used where the data is aggregated and processed to make a decision about the status of the pump. This proposed system ensures that the agricultural farm receives the exact amount of water it needs and prevents water waste. As the moisture content in the soil drops below the specified value, the system will automatically start the pump. Detected parameters and current pump status are displayed on the user's Android application. The farmer uses the Android app to receive updates on the status of his field.

Keywords Cloud computing · Internet of Things (IoT) · MQTT · Soil moisture sensor · ThingSpeak

1 Introduction

Automatic irrigation systems assist farmers in watering crop areas. In the paper [1, 2], a remote irrigation system based on Arduino was developed for agricultural

A. Laha · B. Saha · D. Mukherjee
Dr B. C. Roy Engineering College, Durgapur, India

A. Banerjee
Academy of Scientific and Innovative Research, Ghaziabad, India

P. Karmakar · A. Mukherjee (✉)
CSIR-Central Mechanical Engineering Research Institute, Durgapur, India
e-mail: a_mukherjee@cmeri.res.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_58

plants in remote locations that require water to plant when soil moisture falls below a predetermined threshold. Level of water and the motion of the gate motor that modulates flowrate are sensed in actual time by sensors, which convey the data to the base station or control centre for analysis. Because irrigation is a necessary and fundamental necessity in agriculture, [3] focuses primarily on an efficient automatic irrigation system to save water. Using an automatic irrigation system can help us conserve valuable water resources as well as labour. Because of the significant growth in demand for freshwater, automation technology and its associated equipment, such as solar power, irrigation drip, sensors, and remote control, have become more important in ensuring the best use of water resources. In [4], irrigation is carried out using soil moisture values. To manage the amount of water, the threshold value of a soil moisture sensor is programmed into a microcontroller. In [5], the web application displays the soil moisture values. In [6–8], using the message queuing telemetry transport MQTT protocol, data is forwarded to the cloud and was designed using an industrial IoT architecture with a microcontroller for a device with cloud communication.

In this paper, the proposed system is intended to prevent superfluous water flow into agricultural fields. An IoT-based wireless sensor network is designed and developed, where the soil moisture readings are collected using soil moisture sensors and sent these values to the ThingSpeak cloud. When the soil moisture levels surpass a certain threshold, the relay attached to the controller node operates the pump. A straightforward menu-driven Android application with four settings is developed. This contains information such as pump status and soil moisture level. The motor status indicates the pump's current state. This type of system helps farmers to monitor their agricultural field in real time from remote places and also saves their time, labour cost, reduces wastage of power, and water, with enhancement in crop production.

2 System Architecture

The proposed automatic irrigation system consists of four sensor nodes and one controller node placed in the agricultural field of approximately 100 m^2. The sensor node contains a microcontroller NodeMCU ESP8266 and soil moisture sensor, and the controller node comprises NodeMCU ESP8266 and a relay to control the on–off state of the pump. The soil moisture sensor that is connected to the sensor node, sensed values of the soil parameters from the agricultural field then the sensed data is sent to ThingSpeak and the mobile application as seen in Fig. 1.

The below schematic representation of an Internet of Things-enabled automated watering system. Farmers begin to utilise monitoring different agricultural parameters and control agricultural processes through automation to increase the production of crops.

This proposed work incorporates an embedded system for autonomous irrigation control. The circuit design of one sensor node with a soil moisture sensor is shown in Fig. 2, and the controller node, which runs the pump remotely depending on the

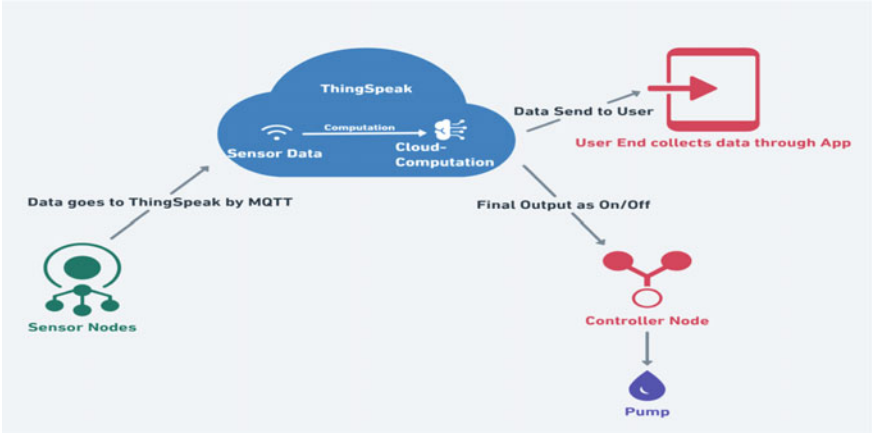


Fig. 1 Schematic diagram of IoT-based automatic irrigation scheduling

threshold value, is shown in Fig. 3. This project includes a wireless sensor network for the real-time irrigation system. This technology ensures a consistent and needed level of water for the agricultural land while reducing water waste. If the soil moisture level is above a certain threshold, the system automatically activates the pump and automatically turns off as the water level returns to normal. The detected parameters as well as the motor's current condition will be presented on the user's Android application.

Figure 4 is an overall block diagram of an IoT-based automatic irrigation system, which consists of four sensors that are connected to the sensor node and provide detected data to ThingSpeak and the mobile application.

The primary goal of this proposed work is to develop an automated irrigation system, which will save the farmer time, money, and energy. Traditional farm land irrigation systems need human involvement. Human intervention can be reduced using irrigation equipment that is automated (Fig. 5).

Fig. 2 Sensor node

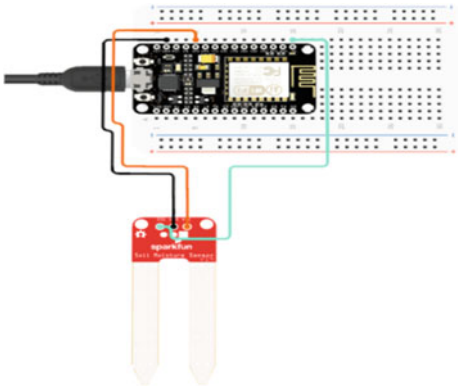
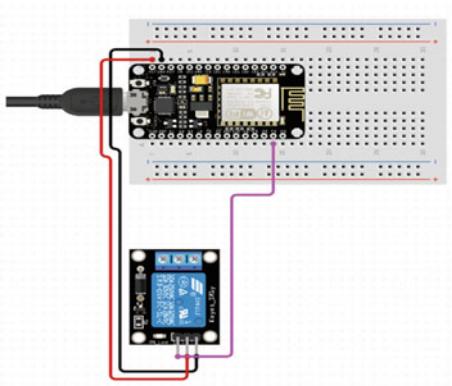


Fig. 3 Controller mode



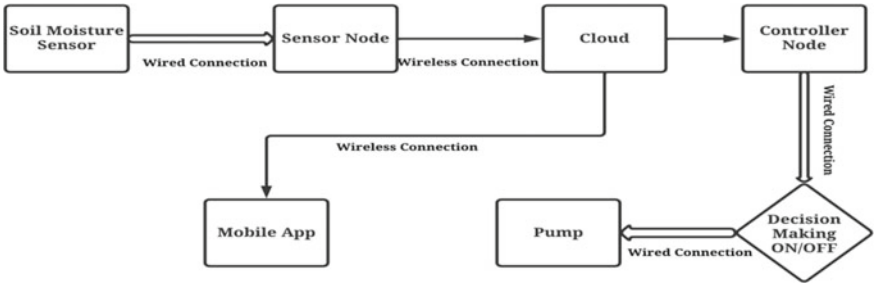


Fig. 4 Block diagram of IoT-based automatic irrigation scheduling

3 Communication Protocol

3.1 MQTT

MQTT is an abbreviation for message queuing telemetry transport. MQTT is an IoT networking protocol that connects machines to machines. It is a publish-subscribe message transport technology that is incredibly lightweight. This protocol is useful for connecting to a remote site where bandwidth is limited. These properties make it helpful in a variety of circumstances, including a consistent environment, such as machine-to-machine communication and Internet of Things contexts. It is a publish and subscribe system in which we may both send and receive messages as clients. It facilitates communication between numerous devices. It is a simple message protocol developed for low-bandwidth devices, making it an ideal choice for the Internet of Things applications.

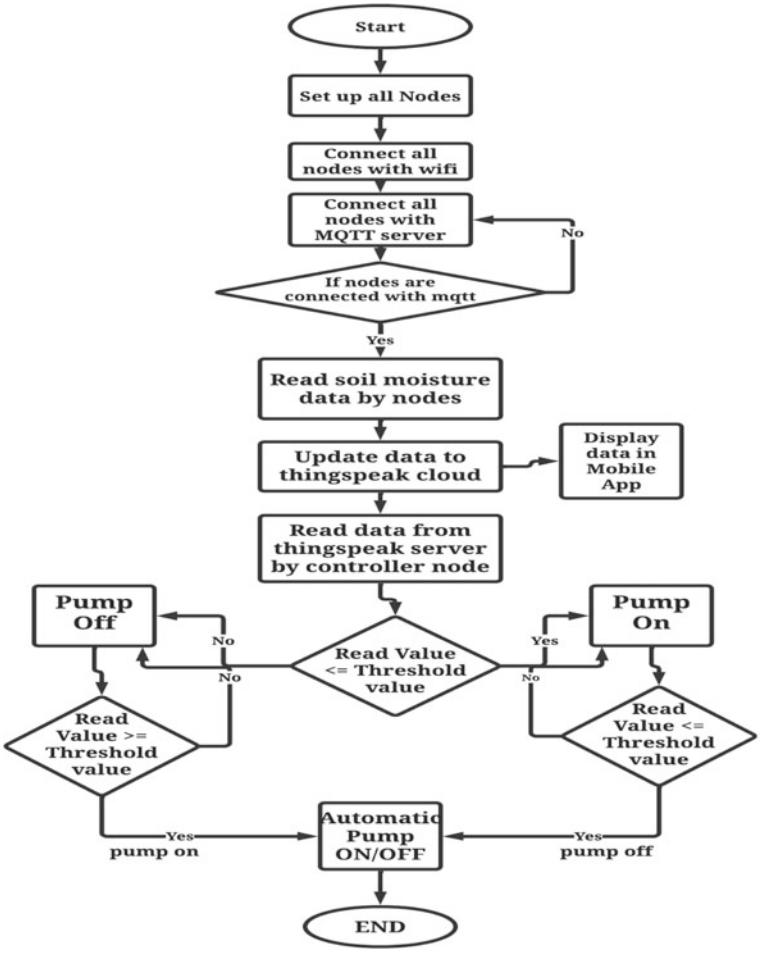


Fig. 5 Flow chart of proposed system architecture

Architecture

To communicate data throughout the architecture, the proposed system uses the ThingSpeak MQTT protocol. This is based on the publish/subscribe approach, in which data from the sensor node acquired by the soil moisture sensor is transferred to the topic soil moisture in the ThingSpeak MQTT server. The data is then sent by the ThingSpeak MQTT server to its clients, such as mobile devices or a backend system. ThingSpeak MQTT Server functions as a broker between the publisher and the subscriber. The publish-subscribe approach is distinct from the client–server model. It distinguishes between the client (publisher) who sends the message and the client (subscriber) who receives the message (Fig. 6).

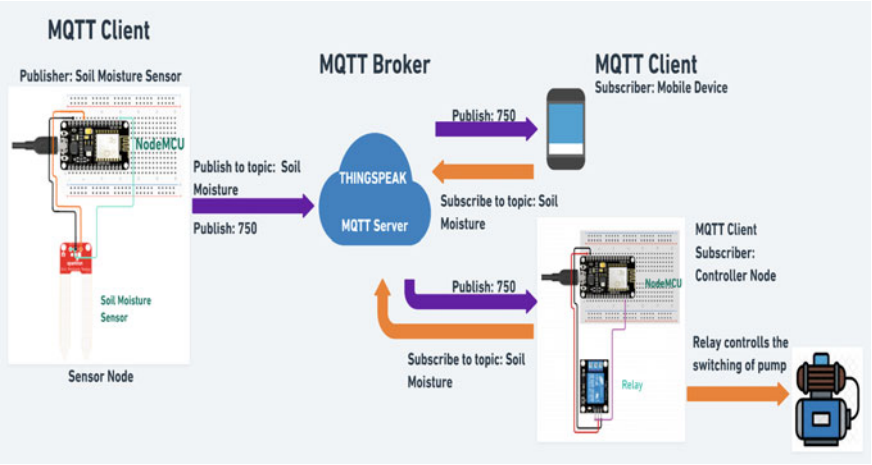


Fig. 6 MQTT publish and subscribe architecture

Following the formation of the Wi-Fi connection, the MQTT connection method begins to construct the connection between the client and the broker in order to publish and subscribe data. First, the connection settings and the PubSubClient instance is setup. The connection is established on port 1883 using the TCP connection type. In the setup procedure, serial data transfer and configuration of the MQTT broker are started. In the loop method of MQTT connection, the MQTT connection is kept open and data to the channel at regular intervals is published. If the connection is lost, the loop method will re-establish it. The ThingSpeak MQTT broker returns with a connectivity response, CONNACK, when the MQTT connection is completed. The MQTT broker delivers a SUBACK message in return to a successful subscription request, and any recent data published to the subscribed channel or field is delivered to the clients.

4 Data Processing, Decision-Making, and Monitoring

The algorithm which has been applied checks the threshold level of moisture content in the soil, if it is greater than the threshold level, it switches on the pump, as it indicates less moisture content. The four sensor nodes collect data from the soil and send it to the ThingSpeak cloud, from there the data is sent to the controller node that is responsible for switching the pump.

Flow chart

See Fig. 7.

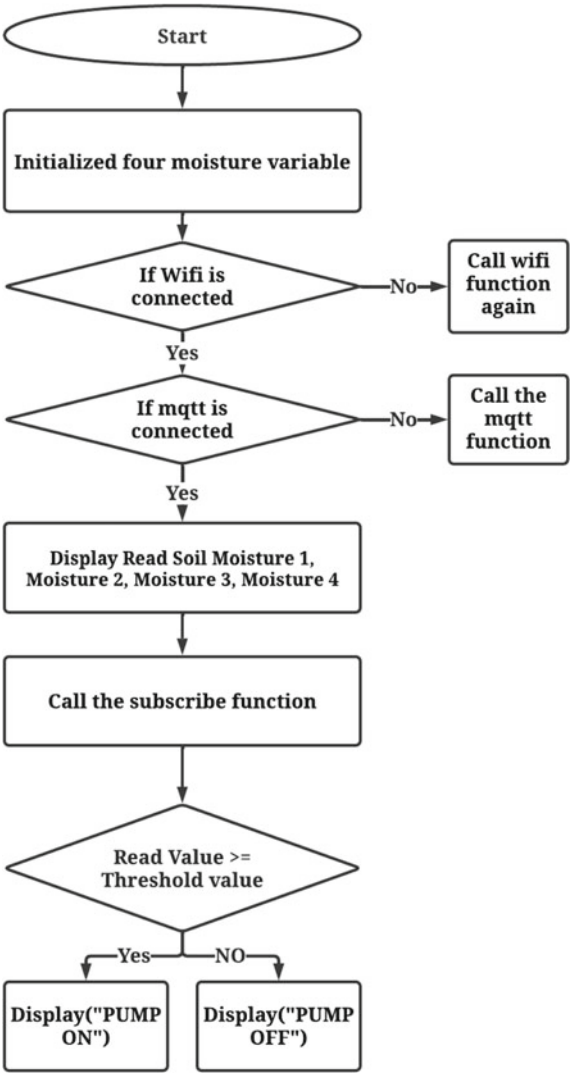


Fig. 7 Data processing and decision-making flow chart

Application Development

The farm irrigation system (FIS) is an Android-based application that will be used by users to check the status of their agricultural field as well as the pump. Through the application, the user will get to know when their field has been watered. At first, the user needs to register himself/herself to the application, after registration, the user could log in to the application. The homepage appears, where the given options are monitoring, documents, feedback, and logout. To see the status of the field and pump, user needs to go to the monitoring page, where the user can check the soil moisture data. To check a particular field like soil moisture, as shown in Fig. 12,

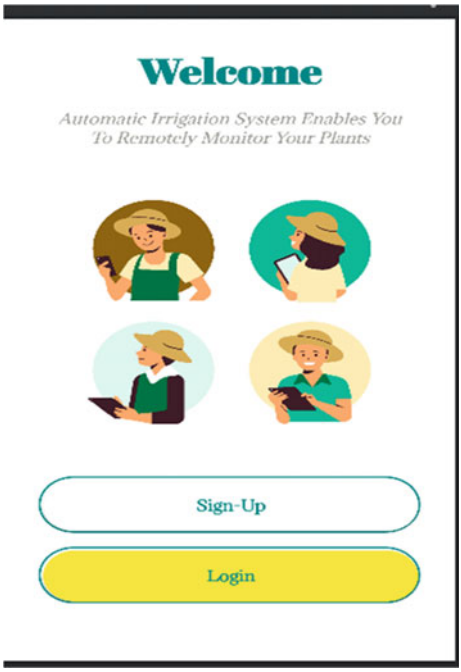


Fig. 8 Welcome screen

the user will click on that and can see the value as well as the graphical plot. The documents section on the homepage will have information about the type of crop and its moisture content. The feedback section is where users will give their valuable responses on the usage of applications. At last, the user will get out from the login page with the logout option (Figs. 8, 9, 10, 11, 12, and 13).

Figures 14 and 15 are graphical plot representations on ThingSpeak. The soil moisture data representation is shown with respect to time, collected from sensor nodes.

5 Conclusion

Agricultural monitoring is critical in order to decrease human labour while also reducing water use. Here, a smart system has been designed that uses wireless sensors to monitor and anticipate soil conditions in order to irrigate the area. Agriculture powered by IoT and contemporary technology has served as a bridge to keep the gap between output and quality from widening. Proper use of IoT will assist farmers with agricultural challenges as well as the rapid growth of technology involved in the agricultural business. Instead of human involvement, the suggested approach involves monitoring crop irrigation by an IoT-enabled agricultural field which delivers precise information to farmers anywhere in the world.

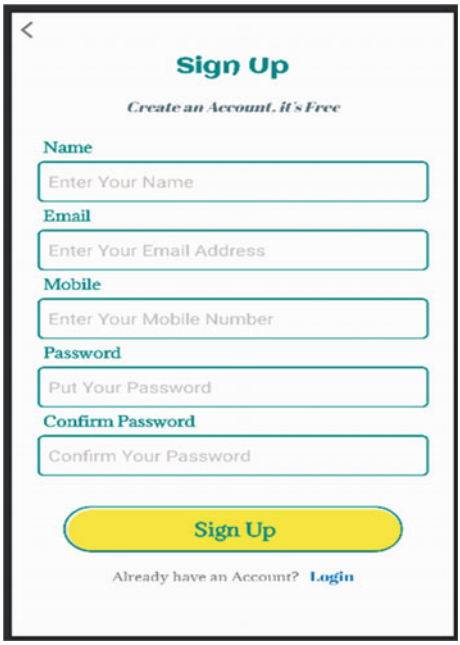


Fig. 9 SignUp or registration screen

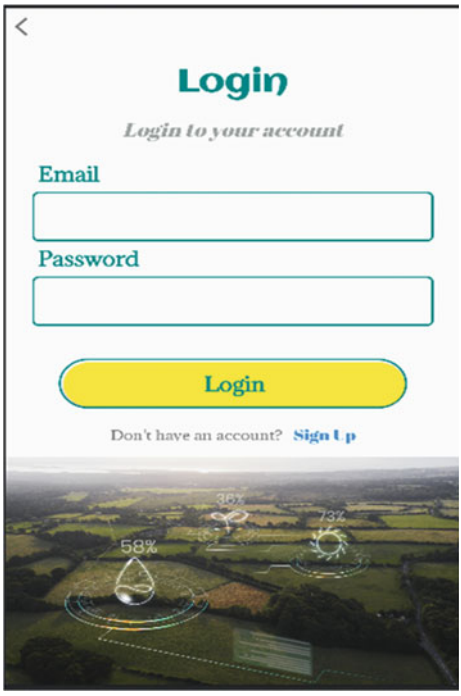


Fig. 10 Login screen

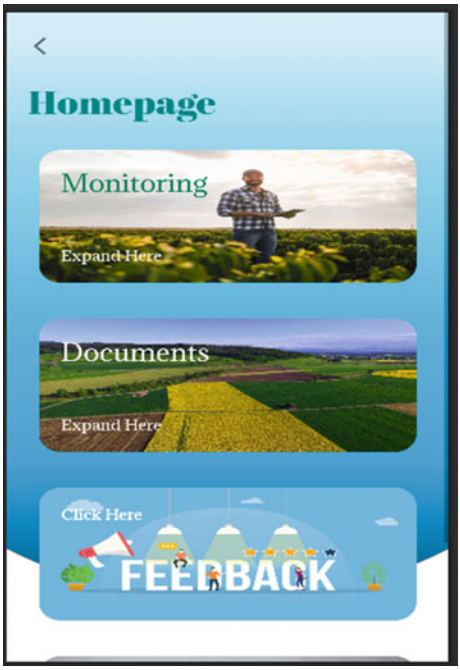


Fig. 11 Homepage screen

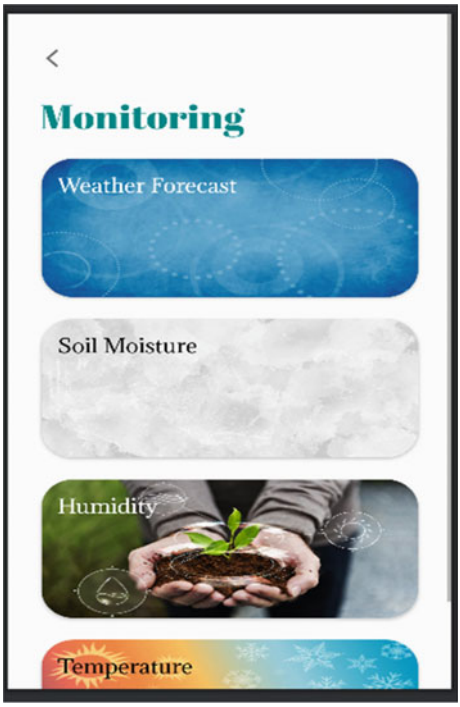


Fig. 12 Monitoring screen

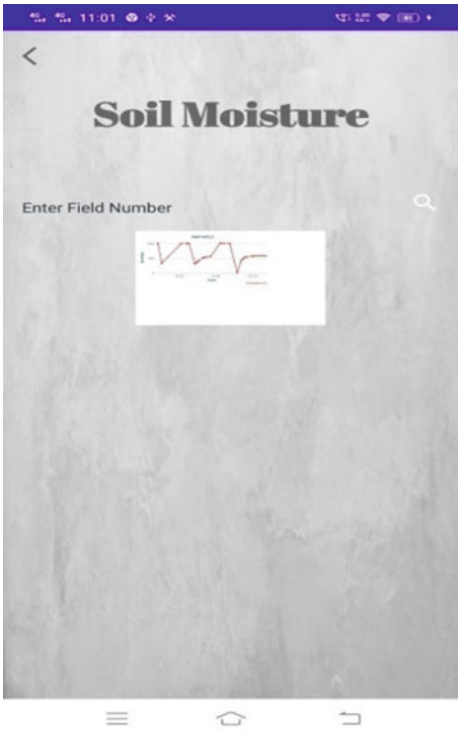


Fig. 13 Soil moisture screen

Fig. 14 Graphical representation of soil moisture data

A customary water system is one of the foremost imperative assignments of cultivating, but these days an IoT-empowered water system framework can give inaccessible access to the agriculturists without going into the field. Proceeded and fast

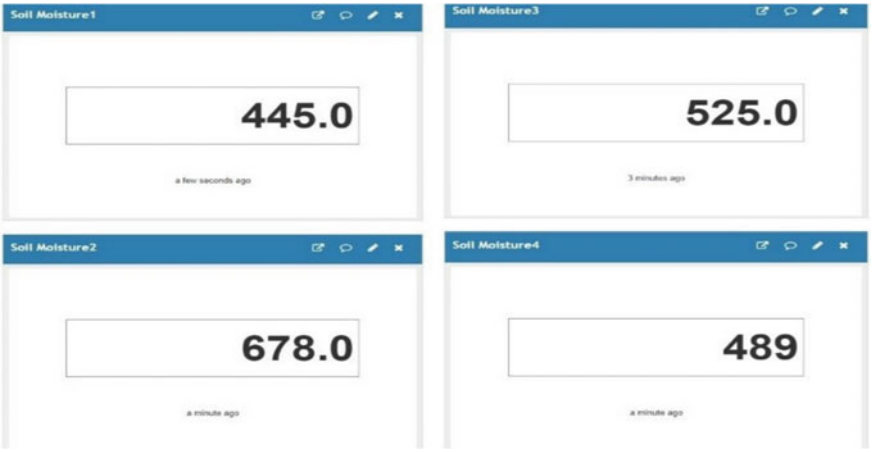


Fig. 15 Soil moisture value from sensor node

advancement of data innovation in agribusiness is an opportunity for trade proficient to investigate and give an effortlessly open and solid framework for savvy horticulture.

Acknowledgements This research is supported by the "Department of Science and Technology", Government of India.

References

1. Hwang J, Shin C, Yoe H (2010) Study on an agricultural environment monitoring server system using wireless sensor networks. Sensors 10(12):11189–11211
2. Arun C, Lakshmi Sudha K (2012) Agricultural management using wireless sensor networks—a survey. In: 2nd international conference on environment science and biotechnology (IPCBEE), vol 48. IACSIT Press, Singapore
3. Shabadi L, Patil N (2014) Irrigation control system using android and GSM for efficient use of water and power. Int J Adv Res Comput Sci Softw Eng 4(7)
4. Kashif Z, Poonguzhali E (2015) IOT based smart irrigation. Int J Eng Res Technol (IJERT) ICESMART—2015 3(19)
5. Anil G, Balakarthikeyan RS, Manikandan S, Prakash S, Varun B (2022) Automatic water saving irrigation system using IoT. Int J Res Appl Sci Eng Technol 10(2):1157–1160
6. Jerald F, Anand M, Deepika N (2019) Design of an industrial IOT architecture based on MQTT protocol for end device to cloud communication. Int J Recent Technol Eng (IJRTE) 7(6S2). ISSN: 2277-3878
7. Kodali RK, Mahesh KS (2016) A low-cost implementation of MQTT using ESP8266. In: 2nd international conference on contemporary computing and informatics (IC3I), pp 404–408
8. Nettikadan D, Raj MSS (2018) Smart community monitoring system using Thingspeak IoT platform. Int J Appl Eng Res 13:13402–13408

Recognition of Struck Out Words Using a Deep Learning Approach

Varsha Naik, Ahbaz Memon, Abhishek Chebolu, Prajakta Chaudhari, and Snehalraj Chugh

Abstract Handwritten document contains a lot of unreadable texts or elements which have no meaning such as struck out words or characters. If such words are fed into a handwriting recognition system, there can be a drop in the accuracy of the system, and it may result into predicting false words. In this paper, we propose an approach to detect these struck out words. With the help of CNN model, we train the model to recognize and differentiate the normal words from the struck out words. For this purpose, some common types of struck out strokes were handled. In order to obtain a handwritten text free of these words, this method may identify strike-through text, locate the word/character, and erase these words. The model was trained on a set of English words and characters that we generated, and it was then put to the test on a range of texts that contained words that had been struck out. The experimental results demonstrate the accuracy of the proposed approach, where the models achieved accuracy levels of 100%.

Keywords Optical character recognition (OCR) · Image classification · Convolution neural network · Struck out words · Handwritten document

1 Introduction

Handwriting recognition has caught the interest of researchers for quite some time. The development of innovative handwriting recognition systems has been made possible by technological advancements in computer architecture, as well as other scientific research developments. Optical character recognition (OCR) is a technique for identifying and recognizing text in images [1]. This process converts the textual data contained in the image into a machine-readable format. It converts the pixel representation of a letter into the character representation that corresponds to that letter. It is a machine that simulates the act of reading. Because of this, OCR of printed text is considered to be a less difficult hurdle to overcome than OCR of handwritten

V. Naik (✉) · A. Memon · A. Chebolu · P. Chaudhari · S. Chugh
Dr. Vishwanath Karad MIT World Peace University, Survey No: 124, Paud Rd, Kothrud, Pune, Maharashtra 411038, India
e-mail: varsha.powar@mitwpu.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_59

text [2] because of the fact that the process can encounter a range of challenges in order to achieve the desired results. The handwritten texts are bound to contain errors or unnecessary elements that can hinder the model's outcomes [3]. The most common of these errors in handwriting is the part of the text that has been struck out or scratched out of the text. When we humans write a word or some content on a page by mistake, we normally scratch or strike-through it to delete it from the page. Due to the fact that it is dependent on the person who wrote the text, every person can strike out in a different ways, such as using crossed-out lines, wavy lines, completely blacked out words, single lines, multiple lines, slanted lines, zig-zag lines, and so on [4]. According to the researchers, the accurate detection of struck out words will aid in a better understanding of handwritten documents [5]. Other types of errors can be reduced by using an image processing techniques. In order to avoid erroneous output, it is essential that the struck out words are handled with care. The paper is divided into five major sections, starting with the literature survey section where we have studied various papers with the implementation of processing of struck out words. Then, in the methodology section, we describe our process and model for the recognition of the struck out words. The numerous applications of the detected words have been described in the applications section of the paper. In the results section, we have noted down the results we obtained from our classification model. The paper concludes with the discussion of the future scope of our study in our conclusion section.

2 Literature Survey

Numerous papers on struck out text processing have been published in recent years [6] has proposed a graph-based model for the representation of a textually connected component in the form of a graph. If the shortest path exists which is almost as long as the text component's width and maintains a reasonable degree of straightness, it is designated as the strike-through line. The accuracy obtained is 98.2% when determining whether a component is strike through. This model dealt with documents written in English, Bengali, and Devanagari scripts.

In the paper [7], the author proposed the use of a two-class HMM-based classifier to distinguish between the words that were struck out and the rest of the words. They examined common properties and characteristics of noisy writing in this paper, using Bangla online handwritten words. Adak and Chaudhuri [6] reports an overall F-measure of 82.02% for identifying strike-through strokes. However, that work detects only strike-through straight lines, whereas [6] considers all types of overwriting, crossing out, and repetition detection, obtaining an F-measure of 71.48%.

The paper [5] has analyzed the effects of strike-through text on writer identification. Writer identification can be viewed as a classification problem in which the objective is to identify a writer from a set of n writers given a sample of handwriting. The paper demonstrates that the presence of struck out texts impairs the ability to identify the writer. For detection, a CNN-SVM hybrid model is used. The results

indicate that they achieved a 98.85% (98.54%) F-Measure for struck out text detection on English (Bengali) handwriting while testing on their generated database. It was observed that the F-measure dropped in the presence of struck out words by 5.43% (4.92%).

In his proposal [8], Alex Shonenkov mentions one of the applications that can be used to strike-through words on handwritten pages. They have proposed a model that generates a person's handwriting from character images that are present in the training dataset. They have attempted to improve the accuracy of handwritten text recognition models by data augmentation. Their proposed system not only generates handwritten text but also stimulates the strike-through characters to the greatest extent possible, making them as close to the originals as possible. In terms of drawing lines that strike out the characters, it has the ability to alter their inclination, size, and transparency.

3 Methodology

3.1 Dataset

In order to carry out the experiment, we required a database of handwritten documents that contained text that had been struck out [9]. As an alternative to using a publicly available dataset, we decided to go ahead and generate our own dataset because we needed a sufficient number of handwritten samples from writers for both training and testing purposes for this research. Consequently, it was necessary to create a new database containing various types of struck-through text in order to deal with the situation. It was requested of the writer that he or she strikes out some words in their running handwriting styles.

There was a total of 500 words collected for each of the two classes, struck out words and normal words, where the struck out words can be seen in Fig. 1. We also increased the dataset size by augmenting the data. When data are augmented, it is advantageous because it improves the performance and outcomes of models by generating new and different examples for train datasets. For each class, a total of 200 images were augmented. The total number of struck out words and the total number of normal words were 700 and 700, respectively.

3.2 Convolutional Neural Network [CNN]

Convolutional neural networks (CNNs), abbreviated as ConvNets, have a deep feed-forward architecture and a remarkable ability to generalize more than networks with fully connected layers [10]. An important distinction between CNNs and ANNs is that CNNs have been extensively used in the field of image recognition, whereas

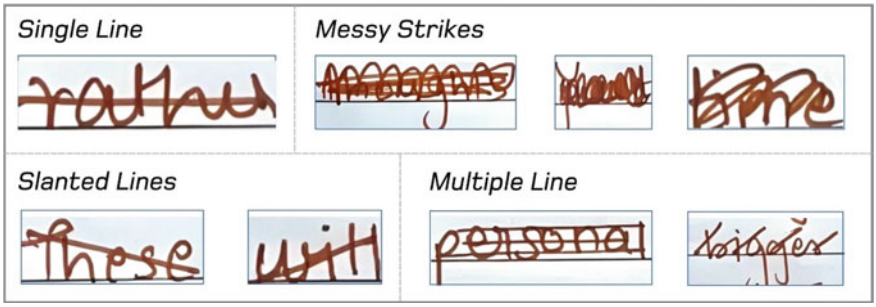


Fig. 1 Different types of struck out words handled in our dataset

ANNs have been used in the field of textual data [11]. The ability to encode image-specific features into the architecture makes the CNN more suited for image focused tasks while also further reducing the number of parameters required to set up the model is a significant advantage.

CNNs are fundamentally based on the presumption that the input will be composed of images. In general, there are three types of layers in CNN's architecture. Convolutional, pooling, and fully connected layers are the types of layers used in CNN Fig. 2. The convolution layers are a collection of convolutional filters or kernels that are used to create the final output [10]. These filters are applied to the input image, which is expressed as N-dimensional metrics, in order to generate the output feature map. The goal of the pooling layer is to reduce the dimensionality of the feature map as it is built up. Each CNN architecture has a fully connected layer at the end of it. Each neuron in this layer is connected to every neuron in the layer before it, forming a network. It is used as the CNN classifier, among other things. Following flattening, this input is in the form of a vector, which is generated from the feature maps that were used to create it. The output of the FC layer corresponds to the final CNN output, which is in the form of a vector created from the feature maps after flattening. The output of the FC layer represents the final CNN output.

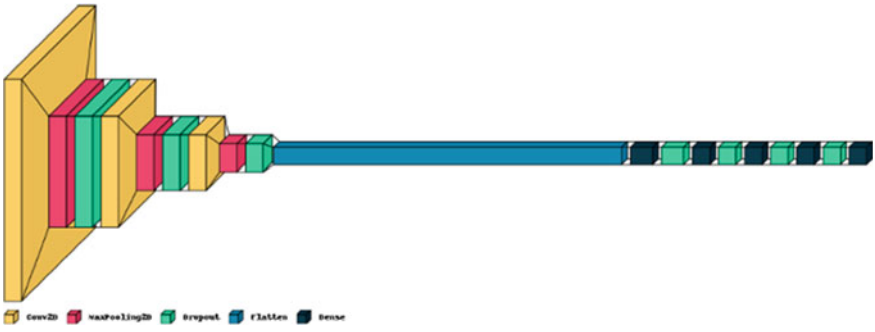


Fig. 2 CNN model architecture

Our dataset consists of images with words and characters, so it was essential to choose a model that would help in handling the data and provide better results. As we discussed, CNNs provide us with an advantage when the dataset comprises of images. We experimented and narrowed down our model for efficient training with the help of CNN architecture. For the training phase, words are extracted from the pages of various writers and label them as normal or struck out. These samples were provided as input to our CNN-based binary classification model [3]. The model is saved so that it can be tested in the future and used for real-time applications. During the testing phase, the extracted words are provided as input to the saved model, which is then used to predict and distinguish the normal words from the crossed-out words. The words that have been crossed-out can either be deleted or used in other applications in the future. This will be discussed extensively in the application section.

4 Applications

Even though the struck out words appear to be meaningless, they can help us achieve a variety of results [12]. In addition to removing the struck out words, our model improves the accuracy of our primary task, handwritten character recognition. After addressing the primary issue of accuracy loss, we can now discuss the potential applications of these struck-through words and how they can help us in various ways.

These words can be used in many ways such as

- They can be used to investigate the patterns and types of strokes used by the writer in order to gain a better understanding of his or her behavior. These cancelled out words or different types of strokes [13] essentially convey the thought process that a writer went through while writing the page in question.
- Using these words, we can figure out which word the author removed or deleted from his or her work. Obtaining this can be accomplished by unmasking the struck out lines from the word and recognizing the characters concealed behind the strike line [14]. A lot of questions, such as which words were misspelled and was the author writing something completely different from the context of his previous words, can be answered.
- Automatically generating handwritten text images is a currently prevalent discussion. The models used for this purpose attempt to replicate human handwriting [15] in order to make it appear more human-like rather than computer generated. It makes use of a dataset that contains the characters and words of an existing author. Because it is important for the generated text to appear genuine and not machine simulated, it is also possible to generate the strike-through words to make it appear more realistic.

Table 1 Classification report

	Precision	Recall	F1-Score
Struck out words	1.00	1.00	1.00
Non-struck out words	1.00	1.00	1.00
Accuracy			1.00
Macro average	1.00	1.00	1.00
Weighted sverage	1.00	1.00	1.00

5 Results

This paper examines the classification of words that have been struck out using a deep learning approach. The CNN architecture is being used as a deep learning model in this study. In order to achieve this accuracy, we trained the CNN architecture from scratch on the dataset we generated. We performed resizing on all of the images to the same height and width because CNN accepts images of the same size. We used 80% of our generated database for training purposes and 20% for testing purposes Table 1. The Adam optimizer was used to train the CNN, and sparse categorical cross-entropy was used as the loss function. Accuracy, recall, the f1-score, and precision are some of the commonly used performance measures that we took into consideration when evaluating the performance of our approach. The previously mentioned model and dataset were used to obtain the results of the above measures. The figure shows that we were successful in achieving a precision and recall of 100% Table 1.

6 Conclusion and Future Scope

In this paper, we discussed the automatic detection of struck out words from English handwriting. These words present a challenge for handwritten text recognition, especially when dealing with complex texts we discussed above, this approach can identify strike-through texts and delete the strike to produce a text that is free of errors. We have tested our model on the English language. Using our CNN model, we were able to achieve an excellent accuracy of 1.0.

In the future, we want to focus on putting the applications into practical methods. If we understand why the writer has crossed-out the word and how the writer has crossed-out the word, the applications listed can provide us with valuable insight into the writer of the document. We can deduce the person's state of mind, while they are writing the page by examining these characteristics. Aside from that, we want to broaden the scope of our model by allowing it to learn on a variety of different languages since each language has its own set of characteristics.

References

1. Smith R (2007) An overview of the Tesseract OCR engine. In: Ninth international conference on document analysis and recognition, vol 2, pp 629–633
2. Namysl M, Konya I (2019) Efficient, lexicon-free OCR using deep learning. In: 2019 international conference on document analysis and recognition (ICDAR), pp 295–301
3. Nisa H, Thom JA, Ciesielski V, Tennakoon R (2019) A deep learning approach to handwritten text recognition in the presence of struck-out text. In: 2019 international conference on image and vision computing New Zealand (IVCNZ), pp 1–6
4. Wang Y, Xiao W, Li S (2021) Offline handwritten text recognition using deep learning: a review. J Phys Conf Ser 1848:12015–12015
5. Adak C, Chaudhuri BB, Blumenstein M (2017) Impact of struck-out text on writer identification. In: 2017 international joint conference on neural networks (IJCNN), pp 1465–1471
6. Adak C, Chaudhuri BB (2014) An approach of strike-through text identification from handwritten documents. In: 14th international conference on frontiers in handwriting recognition, pp 643–648
7. Bhattacharya N, Frinken V, Pal U, Roy PP (2015) Overwriting repetition and crossing-out detection in online handwritten text. In: 3rd IAPR Asian conference on pattern recognition (ACPR), pp 680–684
8. Shonenkov A, Karachev D, Novopoltsev M, Potanin M, Dimitrov D, Chertok A (2021) Handwritten text generation and strikethrough characters augmentation. arXiv preprint arXiv:2112.07395
9. Khosroshahi SNM, Razavi SN, Sangar AB, Majidzadeh K (2021) Deep neural networks-based offline writer identification using heterogeneous handwriting data: an evaluation via a novel standard dataset. J Ambient Intell Humanized Comput 1–20
10. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: 2017 international conference on engineering and technology (ICET), pp 1–6
11. Nath S, Naskar R (2021) Automated image splicing detection using deep CNN-learned features and ANN-based classifier. Sig Image Video Process 15:1601–1608
12. Raundale P, Maredia H (2021) Analytical study of handwritten character recognition: a deep learning way. In: 2021 international conference on intelligent technologies (CONIT), pp 1–5
13. Lemos N, Shah K, Rade R, Shah D (2018) Personality prediction based on handwriting using machine learning. In: 2018 international conference on computational techniques, electronics and mechanical systems (CTEMS), pp 110–113
14. Urs RD, Chethan HK (2021) A study on identification and cleaning of struck-out words in handwritten documents. In: Data intelligence and cognitive informatics. Springer, pp 87–95
15. Guo H (2021) Self-powered intelligent human-machine interaction for handwriting recognition. Research 2021–2021

IoT-Based Smart Monitoring of Soil Parameters for Agricultural Field

Deep Dutta, Chaitali Mazumder, Aishwarya Banerjee, Pratap Karmakar, Debaprasad Mukherjee, and Arpita Mukherjee

Abstract The demand for food is increasing every day because of the growing population. However, the proportional growth in the production of crops is not happening. This problem can be addressed using information technology in the agricultural process. In this paper, an Internet of Things-based (IoT) system has been designed and developed to monitor soil health. The chemical and physical properties of the soil can be evaluated by measuring different parameters contained in the soil such as nitrogen (N), phosphorus (P), potassium (K), temperature, moisture, and salinity as well as electrical conductivity. Three sensor nodes are developed using NodeMCU ESP8266, NPK sensor, EC sensor, soil moisture, and temperature sensor. First, the data obtained from the procured sensors have been tested and validated in the laboratory by adding different amounts of urea and water. These sensor nodes, which are powered by solar, are placed at some desired location in the agricultural fields to collect the soil data. All the sensor data are then sent to the cloud over a Wi-Fi network using Message Queuing Telemetry Transport (MQTT) protocol for monitoring, further analysis, and decision-making. The data of the soil parameters are stored in a cloud platform, which will allow monitoring of the data in graphical and numerical formats. An android application is developed for real-time monitoring of the data.

Keywords Cloud server · Internet of Things (IoT) · MQTT (Communication protocol) · Soil sensors · Soil nutrients analysis · Soil health monitoring

D. Dutta · C. Mazumder · D. Mukherjee
Dr. B. C. Roy Engineering College, Durgapur, India

A. Banerjee
Academy of Scientific and Innovative Research, Ghaziabad, India

P. Karmakar · A. Mukherjee (✉)
CSIR-Central Mechanical Engineering Research Institute, Durgapur, India
e-mail: a_mukherjee@cmeri.res.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_60

1 Introduction

Agricultural practices are closely associated with the production of essential crops and are the main source of income for most developing countries. Agriculture plays a major role in the economic growth of developing countries like India, Nepal, Pakistan, and Sri Lanka. A nation's food security is ensured by the agricultural sector. Agriculture is the backbone of the country's economic system, about 70% of the people rely on it. There is a high demand for food, but a low supply. In comparison to population growth, agricultural production is insufficient. A primary cause of this decrease in agricultural production is the use of old conventional methods by farmers. Due to continuous changes in the weather, soil conditions, and climatic conditions, and a lack of knowledge and technology, it is very difficult for farmers to keep track of all these records, analyze, and harvest crops appropriately. To address this problem, an IoT-based (Internet of Things) soil health monitoring system is proposed; this will enable users to determine the health of the soil to maximize crop productivity.

A brief background study on sensor-based soil monitoring systems is discussed below. The paper [1] presented a system where the data from the soil sensors are delivered to the MCP3204 ADC (Analog-to-digital converter), which then sends it to the cloud through Raspberry Pi, and the information may be accessed via mobile phones and personal computers (PCs). The study in [2] consists of an electronic device that can detect macronutrients using three light-emitting diodes (LEDs) as a light source and one light-dependent resistor (LDR) as a light detector. Light converts the remaining light reflected by the reflector into an electric current. The system uses a microcontroller to collect the data, so the output of the converter is converted to reading on a digital display. The research in paper [3] helps farmers to receive suggestions for sulfur and lime based on the pH value of the soil. A t-test showed that there was no significant difference between the calculated data and the recorded data in the laboratory, confirming the proposed approach, and from the proposed system agriculture scientists, farmers, IoT experts and agriculture professionals, will be benefited. The proposed system in [4], the soil health monitoring unit (SHMU), is photovoltaic and can be installed in the field for long periods. Using LoRaWAN wireless technology, each SHMU wirelessly sends soil temperature, humidity, conductivity, CO, and geographic location data and uses a dashboard to analyze soil health measurements for soil. The research in [5] describes a soil sensing methodology that uses a nitrogen phosphorus potassium (NPK) testing kit and other traditional methods to measure digitally using a cloud-based prognostic technique that helps increase crop output with minimal water and fertilizer use. The author of the paper [6] used IoT technology to detect and control devices remotely via an access network, allowing numerous physical objects to be directly connected to a computer system. They use temperature sensors, pressure sensors, and humidity sensors, and Lora is used as a wireless communication system, which is a low-power wide area network (LPWAN) technology. The paper [7] attempts to introduce an embedded soil monitoring system that measures soil moisture and performs automatic irrigation, ambient temperature, and humidity. Use the PHP (Hypertext preprocessor) Web site to change the

configuration data for watering timing and soil water content lower and upper limits, and monitor real-time soil water changes in charts. The paper [8] describes a low-cost system that can obtain up-to-date information on crop surroundings and farm issues and proposes the best crop needs based on data from sensors. The proposed system is an IoT-based crop monitoring system using Global System for Mobile communication (GSM) technology. The paper [9] presents an overview of various soil sensors used to measure temperature, moisture, and humidity. Information from the sensors on the ground is displayed in the android app which is very helpful in making cultivation recommendations to farmers for better agriculture.

The proposed IoT-based soil health monitoring system helps to monitor real-time soil moisture content and electrical conductivity of the soil water. It detects the nutrient values of the soil like phosphorus, nitrogen, and potassium. It is low cost and powered by solar systems. The data can be monitored using a smartphone or laptop/desktop.

2 IoT-Based Soil Health Monitoring

An IoT-based soil health monitoring system has been proposed. This helps to monitor various soil parameters and to make appropriate decisions regarding the use of the right amount of water and fertilizer. This monitoring system will also help to maintain all these soil parameters at their desired level. Some soil physicochemical properties, such as water content, temperature, soil nitrogen, phosphorus, and potassium levels, have a significant impact on crop yields. The proposed system consists primarily of various sensors for monitoring soil conditions and NodeMCU ESP8266 for data acquisition and sending it to the cloud for monitoring and further analysis. The solar power bank has been used to power the sensor nodes. Three sensors track and monitor soil moisture, soil nutrient levels, and soil electrical conductivity. The sensors detect the soil parameters data and send it to the Thingspeak cloud platform using the MQTT communication protocol.

3 Hardware Used

It is an IoT-based soil health monitoring system that monitors various soil parameters to improve crop yields. Various sensors are used to determine soil parameters and monitor soil health. Three sensors are used here: EC sensor, soil moisture sensor, and NPK sensor. These sensors are attached to a microcontroller NodeMCU ESP8266, and data are sent from the microcontroller to the cloud server using the MQTT communication protocol.

3.1 NPK Sensor

The soil NPK sensor can detect the levels of nitrogen, phosphorus, and potassium in the soil. It aids in examining soil health quality, enabling a much more systematic assessment of soil condition. A Modbus module is required to read the NPK data. The Modbus module communicates with the microcontroller and the sensor.

3.2 Electric Conductivity Sensor (EC)

EC sensor measures the electrical conductivity of a solution. It is a low-cost system having a satisfactory high accuracy rate of conductivity of electricity. Direct measurement of soil water can be captured in the field using EC water conductivity test equipment. A voltage is applied between the electrodes by the sensor placed into the soil water sample. The resistance of water is calculated from the voltage drop, which is converted to conductivity. Conductivity is inversely proportional to resistance and is expressed as the amount of conductance over a given area. If the EC value is too low, it indicates that the soil is deficient in nutrients, and if the EC value is too high, it indicates that the soil is rich in nutrients.

3.3 Soil Moisture Sensor

It is normally used to estimate the soil water content. This sensor is made up of 2 probes that measure the content of the water. If there is more water, it conducts more electricity, resulting in much less resistance; as a result, the moisture content is higher and vice-versa. Less than 500 is too wet, 500–750 is the target range, and greater than 750 are dry enough to be watered.

3.4 Microcontroller (NodeMCU ESP8266)

NodeMCU ESP8266 is an open-source platform that connects devices and enables transferring of data using Wi-Fi protocol. It is compatible with Arduino IDE and is very easy to use. Multiple sensors are directly connected to it for measuring and monitoring the soil health parameters. NodeMCU along with smart sensors measures the quality (moisture, nutrition, conductivity) and effectiveness of soil. The controller uses sensors to access the monitored information. The measured data will be sent and displayed on the cloud platform. This experimental result monitors the soil quality in real time.

4 System Architecture

The structure, behavior, and viewpoints of a system are defined by the system architecture, which is a conceptual model. An architecture description is a formal specification and representation of a system that is organized in a way that makes thinking about the system's structures and behaviors easier. The block diagrammatic representation of the proposed system is given (Fig. 1). The electronics circuit diagram of a sample sensor node is given in Fig. 2. The three solar-powered sensor nodes installed in agricultural fields are given in Fig. 3.

4.1 Communication Protocol (MQTT)

The proposed model uses the Message Queuing Telemetry Transport (MQTT) protocol to connect various IoT remote devices with the server with minimal network bandwidth. MQTT (in Fig. 4) is an Internet of Things communication protocol that has been certified by Organization for the Advancement of Structured Information Standards (OASIS). It is designed as a very lightweight subscribe/publish messaging

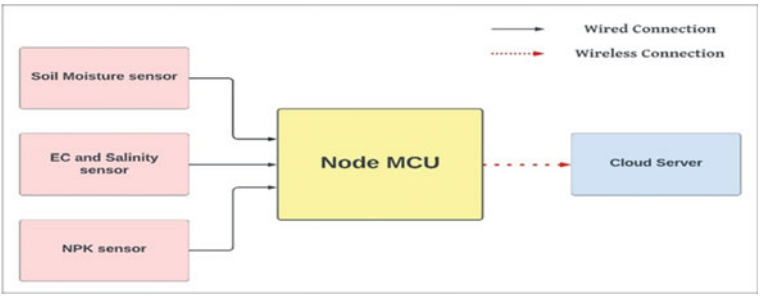


Fig. 1 System block diagram

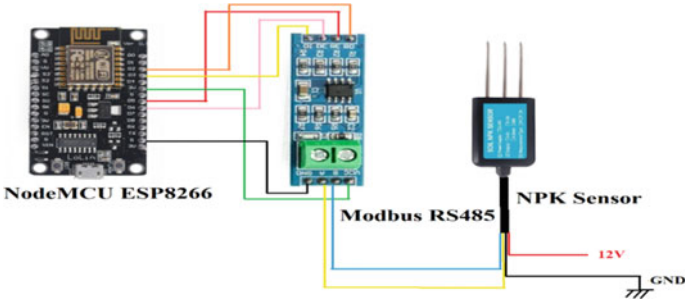


Fig. 2 Circuit diagram of the sensor node system

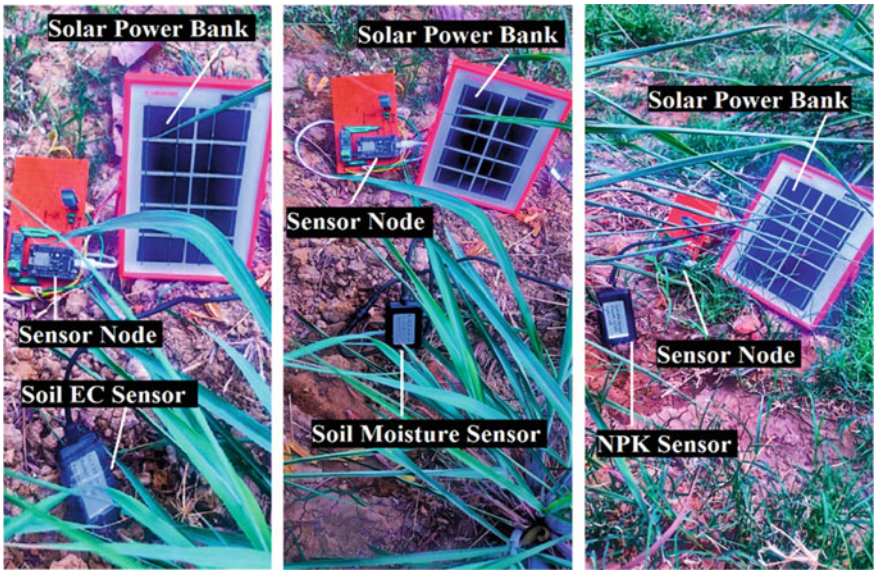


Fig. 3 Connections of the three different solar-powered sensor nodes installed at the agricultural field

transport which is ideal for connecting remote devices. Although MQTT is most commonly used with transmission control protocol/Internet protocol (TCP/IP), it is used with any network protocol that enables lossless, ordered, bidirectional communications. A message broker and clients are the two sorts of network entities defined by the MQTT protocol. The MQTT broker is a server that accepts all messages from clients and forward them to the target client. The MQTT client is a device that runs the MQTT library and connects to the MQTT broker over the network. MQTT clients use the "Topic" subject line instead of the intended recipient's direct connection address. Anyone who subscribes to a topic receives a copy of all messages related to that topic. Transport Layer Security (TLS) encryption is used by MQTT, and connections are protected. The data are collected using sensors connected to a microcontroller and transferred to the cloud using the MQTT communication protocol, which will be routed through an MQTT broker in real time. The information will be sent to a cloud server (ThingSpeak) and can view it all on the PC or mobile device. It is important to note that the sensor will continue to communicate new data to the cloud server without waiting for a response. The Web server will issue a request to retrieve messages/data sent by the sensors. MQTT is ideal in the field of IoT since the entire message sending process is an example of asynchronous communication.

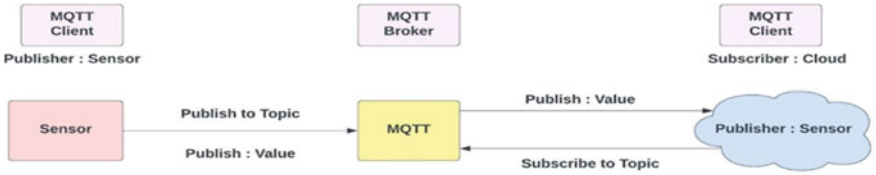


Fig. 4 MQTT protocol

Fig. 5 Client and server interaction diagram

4.2 Client and Server Interaction

The user initiates a request, which is then accepted by the cloud server. The server acts as a channel between the user's request and the server's response. The request is then accepted by the server, and the user is given access to the account. The user may observe, monitor, and analyze the soil's many parameters. A block diagram representation of the client and server interaction is shown in Fig. 5.

5 Monitoring of Data

The user can remotely monitor the various soil parameters from any remote place by using our application "Farmdoc." It is an android application that is used to monitor soil data in real time. This application is built using MIT app inventor. The application consists of a splash screen that displays the app name, and a homepage where the soil data will be displayed in the form of graphs. The soil data are fetched from the cloud server, i.e., the ThingSpeak server, and the application shows all the relevant data. It displays various soil parameters like potassium content, moisture content, phosphorus content, electrical conductivity, nitrogen content, and salinity level of soil in graphical format (in Fig. 7). The flowchart of the proposed system is given (Fig. 6).

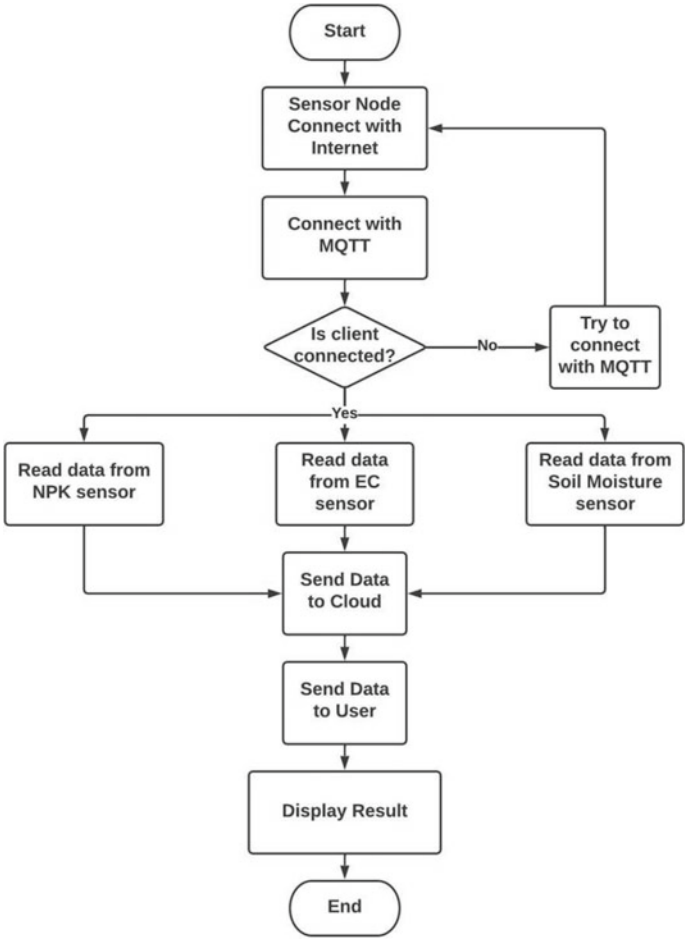


Fig. 6 Flowchart of the health monitoring system

6 Results and Discussion

In this paper, a soil health monitoring system is developed that is used to monitor electrical conductivity, nitrogen content, phosphorus content, potassium content, and soil moisture content of the soil in the agricultural field. Three different sensors are used to monitor these parameters. At first, these sensors are calibrated and validated in a laboratory setup to test the variation in the sensor's reading with changes in soil condition.

One of the most important parameters in determining soil health is NPK values. When the NPK values are low, then urea is added to the soil. If the NPK values are very high, then any particular type of crop may not be cultivated there. The reading obtained from the NPK sensor has been tested by adding varying amounts of urea

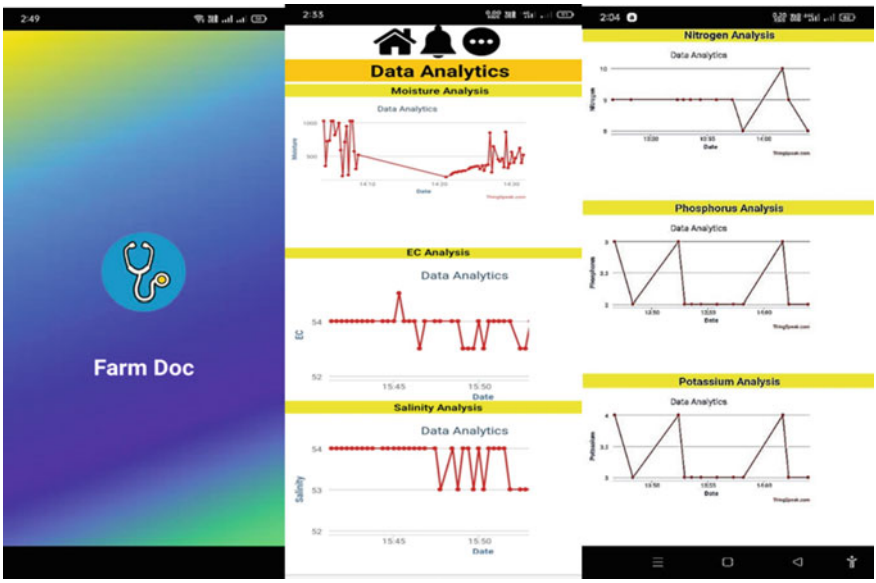


Fig. 7 Various soil property data displayed in the developed application

to a soil sample. 1 kg of a soil sample is taken and divided into two parts—Soil Sample-1 and Soil Sample-2. The nitrogen content of Soil Sample-1, i.e., plain soil without urea using an NPK sensor is measured. Then, 5 pieces of diluted urea are added to Soil Sample-1, and then again, the nitrogen content is measured using the NPK sensor. There is a very slight increase in NPK level. After taking the NPK reading, the percentage difference of the Soil Sample-1 without (n_1) and with (n_2) urea using the formula $((n_1 - n_2)/n_1)$ *100 is calculated. The percentage difference ranges from 2 to 5%. The nitrogen content of Soil Sample-2, i.e., plain soil without urea (n_3) is measured using an NPK sensor. Then, 10 pieces of diluted urea are added to Soil Sample-2 (n_4), and then again, the nitrogen content is measured using the NPK sensor. There is a large increase in NPK levels. After taking the NPK reading, the percentage difference of the Soil Sample-2 without and with urea $((n_3 - n_4)/n_3)$ *100 is calculated. The percentage difference ranges from 38 to 41%. This test can provide an idea of how much urea should be added to the field for a particular type of crop in presence of reference data. The NPK values are recorded and monitored through the mobile application as shown in Fig. 8.

When EC exceeds the desired level, soil fertility decreases. To test the EC and salt content of the soil, a 1 kg soil sample is taken. Electrical conductivity and salinity of soil sample without water are measured using an EC sensor. It is observed that the values for EC and salinity are the same. When 100 ml of water is added to the soil, there is a slight change in the values of EC and salinity. First, the values increase then decrease, and finally, they become equal when the water dries up. Then, 200 ml of water is added to the soil; there is a large change in the values of EC and salinity.

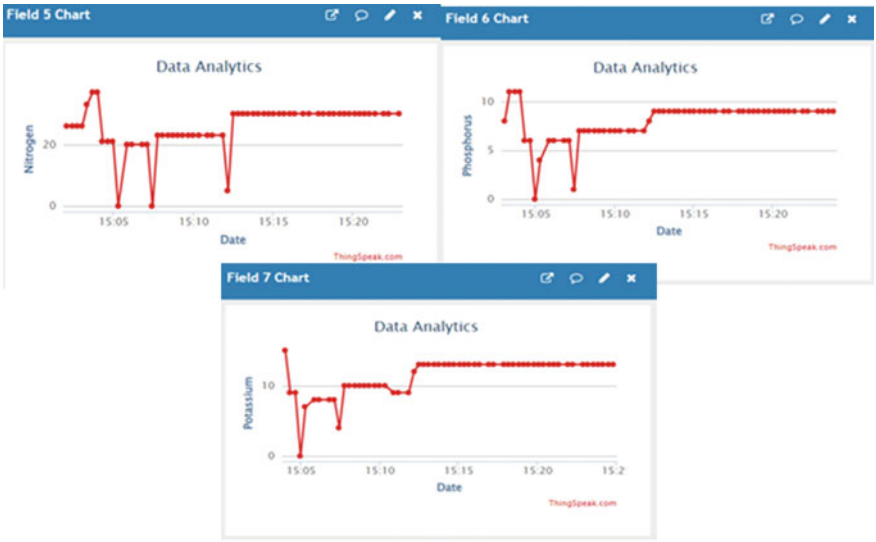


Fig. 8 NPK sensor data displayed on the cloud

First, the values increase then decrease, and finally, they become equal when the water dries up. When the value of electrical conductivity is high, it needs to be reduced, and this can be done with the process of leaching where water needs to be added to the soil so that the amount of salt can be reduced from the soil. It is observed that if less amount of water is added, then the values of EC and salinity are less. If more amount of water is added, then the value will be more. The data of EC and salinity of soil can be monitored from the user device as shown in Fig. 9.

To test the moisture of the soil, a 1 kg soil sample is taken. The soil moisture sensor data are recorded by gradually increasing the amount of added water in the soil sample. It is observed that moisture content increases as the water content of the soil increases. First, adding 50 ml of water will increase the moisture content of the soil, which will become constant after a while. Then, 100 ml of water is added; there is a slight increase in the moisture content of the soil, and after some time, moisture

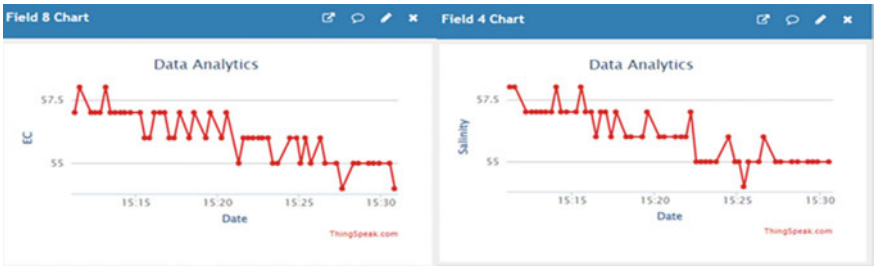


Fig. 9 EC sensor data displayed on the cloud

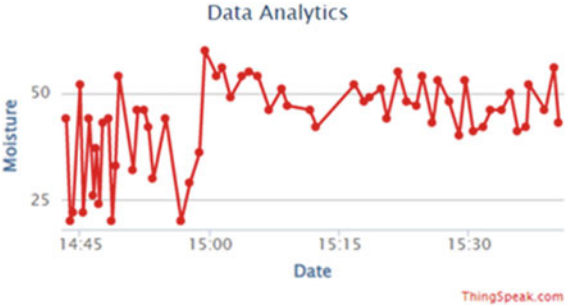


Fig. 10 Soil moisture sensor data displayed on the cloud

content becomes constant. The moisture sensor data as displayed in the user device are shown in Fig. 10. Initial obstacles faced are the issue of power supply in the field which is solved using the solar power.

7 Conclusion

The proposed solar-powered soil health monitoring system is a cost-effective solution to the problem of the farmers. They can access and monitor the soil parameters from anywhere in the world at any time; it makes their lives easier. The data collected by the sensors are continuously sent by NodeMCU microcontroller using MQTT to the cloud server wirelessly in real time so that the farmers can take necessary steps according to the change of soil parameters for better crop yield. It would help the farmer to use the right amount of fertilizer and water depending on the requirement of any particular type of crop. The future scope of this proposed system is to monitor soil parameters in a more detailed way with utmost precision and to provide accurate suggestions after analyzing the soil health report of the particular field. The suggestion would help the farmer to take the necessary steps for better growth of crops at all times irrespective of their knowledge about the agricultural sector.

Acknowledgements This research is supported by the "Department of Science and Technology," Government of India.

References

1. Manickam S (2020) IoT-based soil condition monitoring framework. SSRN Electron J
2. Goswami V (2020) Soil health monitoring system. IJRASET 8(V). ISSN 2321-9653
3. Bhatnagar V (2020) IoT-based soil health monitoring and recommendation system. Int J Innov Implementations Eng 1. ISSN 2454-3489

4. Ramson S, León-salas W, Brecheisen Z, Foster E, Johnston C, Schulze D, Filley T, Rahimi R, Soto M, Bolivar J, Bolivar L, Málaga M (2021) A self-powered, real-time, LoRaWANIoT-based soil health monitoring system. 8(11):9278–9293
5. Patidar P, Joshi S (2019) Design and development of cloud-IoT-based soil health monitoring system. 7(08):144–149
6. Thereza N, Saputra A, Hamdadi A (2020) The design of monitoring system of smart farming based on IoT technology to support operational management of tea plantation. In: SICOINIAN 2019, vol 172, pp 52–57
7. Juca S, Rabelo L, Gonçalves C, Silva F, Pereira S, Da-Silva A (2018) Construction of soil moisture and irrigation IoT monitoring system using Project Based Learning. Int J Innov Educ Res 6(8):99–111
8. Grace M, Mohanan A, Amina A, Antony S, Aryaraj K (2018) Crop health monitoring using GSM technology. In: ERTE 19, pp 108–115. e-ISSN 2455-5703
9. Patil V, Hake B, Mohite V, Pawar S, Hambarde S (2019) Soil monitoring system for crop recommendation. IJSART 5(6):3–7

Dual-Active Bridge Converter with Single Phase-Shift Control for Distribution Solid-State Transformer

Mohmmed Rizwan Ansari and D. K. Palwalia

Abstract Solid-state transformer (SST) is essential equipment in electrical devices. Because of its compact size, it is often employed in high-power applications. Converter-based solid-state transformer has various features such as low weight and size, high frequency, and no core losses. These capabilities make a solid-state transformer replace the power distribution transformer. Bidirectional power flow is also possible in the SST, so it can be easily integrated with the utility grid and renewable energy storage or larger batteries. SST can be made up of three stages: rectifier, dual-active bridge (DAB) converter and inverter. DAB converter is a vital component of SST and is used for power transfer from one circuit to another circuit. Two h-bridges with eight IGBT switches and one high-frequency transformer (HFT) serve as isolation in the proposed DC-DC DAB converter. As an application of SST, a DAB converter was studied and simulated in a MATLAB Simulink environment and phase-shifted pulse generated from dSPACE 1103 platform. This work employs the single-phase shift (SPS) control approach to provide steady output and power transmission.

Keywords Solid-state transformer (SST) · High-voltage DC (HVDC) link · Low-voltage DC (LVDC) link · Single-phase shift (SPS) · And dual-active bridge (DAB) converter

1 Introduction

Solid-state transformer is the main component of the future smart distribution system. As the population increases day by day, non-renewable energy plants such as thermal plants can't deliver the power continuously in the next years. So, the solution is dependency on renewable sources such as solar and wind. But, they cannot be integrated

M. R. Ansari (✉) · D. K. Palwalia
Rajasthan Technical University, Kota, India
e-mail: mohmmedrizwan8@gmail.com

D. K. Palwalia
e-mail: dkpalwalia@rtu.ac.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_61

with the conventional transformer due to their unidirectional property. SST is able to integrate with renewable sources because it has bidirectional power flow. SST has several advantages such as low weight, compact size, compensation of voltage sag and swell, fault isolation, harmonic compensation, and voltage regulation [1]. SST can be used in microgrids, traction systems, and aerospace applications [2–4].

As illustrated in Fig. 1, SST has several topologies depending on stages, including single stage, two stage, and three stage. Single-stage SST has AC to AC conversion in type-A; two-stage SST has two parts: DC link situated at high voltage side in type-B and DC link situated at low voltage side in type-C; three-stage SST has three conversion in type-D where, inverter called input stage, DAB converter called isolation stage and rectifier called output stage, and it has both DC link high voltage as well as low voltage [5]. Three-stage SST provides better controllability and flexibility in the system. SST can be used in a power system especially in distribution system where it can perform all the functions of FACTs and STATCOM devices [6]. Figure 2 depicts a single-phase SST circuit diagram. This SST has an h-bridge inverter, a DAB converter with two h-bridges interlinked with HFT, and an h-bridge inverter.

For isolation stage, where DC to DC conversion is performed, numerous converters are used for this stage such as LLC converter, current doubler, and push–pull converter etc. [7, 8]. Among all of these DAB, converter is widely used due to its simple structure, low cost, and less resonant component [9]. The study of DAB converter using a SPS control approach is addressed in this work, and it was also simulated and evaluated in the MATLAB Simulink.

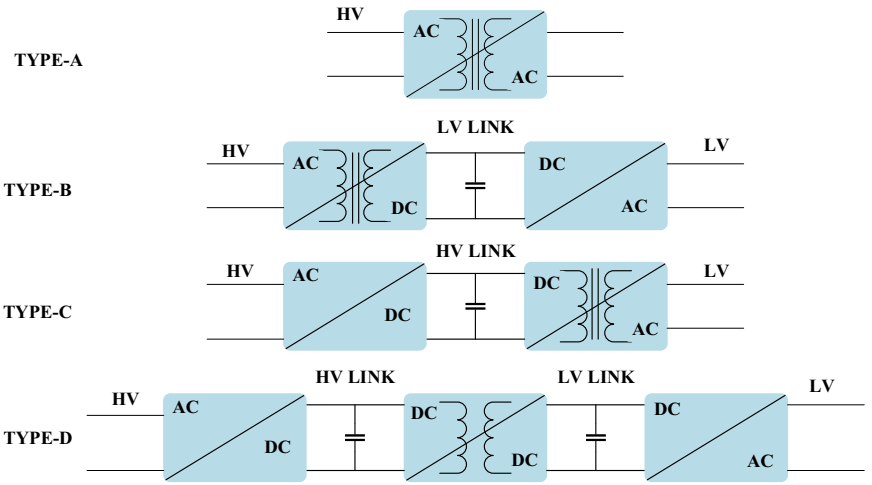


Fig. 1 SST topologies

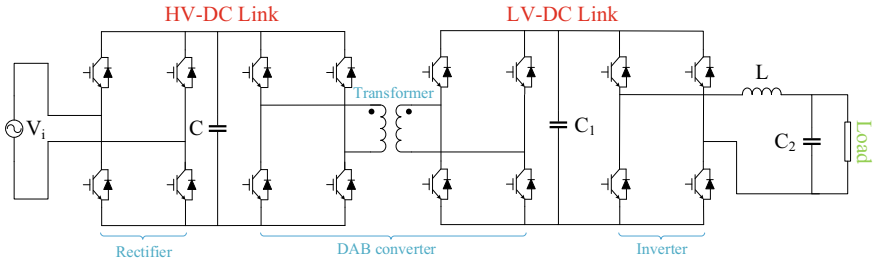


Fig. 2 Single-phase SST circuit diagram

2 History of DAB Converter and Circuit Configuration

Resonant DC-DC converters are used with HFT to reduce the weight and filters. In low-power applications, they provide soft switching during operation with reasonable heat sink size [16–18]. In high-power applications, these converters required additional snubber inductance to transfer the energy and RC networks for soft-switching operation that increases conduction and switching losses and decreases the efficiency of the resonant converters. For high-power applications, the first DAB converter is proposed in 1990 by Doncker and Divan [10]. This converter consisted of two bridges and one HFT. IGBT semiconductor device with anti-parallel diode is used as switches for bridge, and isolation is provided by HFT. The leakage inductance of HFT is used to transfer the energy from one bridge to another bridge. The DAB converter has the ability of bidirectional power flow and low component stress and soft-switching achieved by a capacitor connected in parallel with semiconductor switches during operation [11].

The proposed single-phase DAB converter circuit diagram is depicted in Fig. 3. Each h-bridge is made up of two legs. Four IGBT with four antiparallel diodes has been employed on the LV side. Antiparallel diodes D_1, D_2 and D_3, D_4 have been linked across switches M_1, M_2 and M_3, M_4. Similarly, the switches M_5, M_6 and M_7, M_8 have been linked in antiparallel across the diodes D_5, D_6 and D_7, D_8. Two h-bridges are linked by a high-frequency transformer. The main h_1-bridge is used to apply the input signal V_1, and output voltage V_2 is evaluated at h_2-bridge. Switches M_1 and M_3 are switched on at time t_1, while switches M_5 and M_7 are turned on after a 50% delay. Both switches M_2 and M_4 are both turned on at same time t_2. Switches M_6 and M_8 are turned on after giving the phase-shift α_1.

3 Single-phase Shift (SPS) Control

This control technique is widely used in high-power applications due to its simplicity and ease of control [12–14]. Bidirectional power flow is also possible in SPS control technique. This technique has benefit of small filter, low sensitivity to parasitic

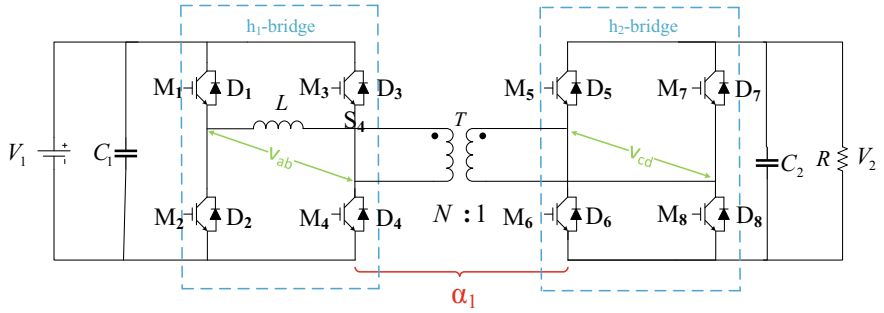


Fig. 3 Circuit diagram of single-phase DAB converter

components, and good efficiency when voltage conversion ratio k is equal to 1 ($k = V_1/N\ V_2$). Standalone SST without control technique gives large inductor current that increases the conduction losses and core losses and decreases its efficiency. These issues are overcome by using SPS control technique [15].

In SPS control strategy, a 50% duty ratio of switching time T_s is provided to M_1 switch, and complementary of M_1 gate pulse is provided to M_2 switch. Remaining 50% duty ratio of T_s is given to M_3 switch, and complementary goes to M_4 switch. Some phase-shift ratio α_1 is provided to M_5 switch from switch M_1 pulse with the help of closed control which is 50% of T_s and complementary goes to M_8 switch. Remaining 50% duty ratio of T_s is given to M_6 switch, and complementary goes to M_7 switch. By applying these pulses, the voltage of h_1-bridge v_{ab} and voltage of h_2-bridge v_{cd} are square and α_1 phase shifted to each other as shown in Fig. 4. This α_1 helps to transfer the power from h_1-bridge to h_2-bridge.

The transferred power in DAB converter for SPS control from one bridge to another bridge is given as:

$$P = \frac{NT_sV_1V_2}{2L}\alpha_1(1-\alpha_1) \tag{1}$$

where

V_1 = input voltage
V_2 = output voltage
N = turn ratio of HFT
T_s = switching time
N = turn ratio
L = total leakage inductance of HFT.

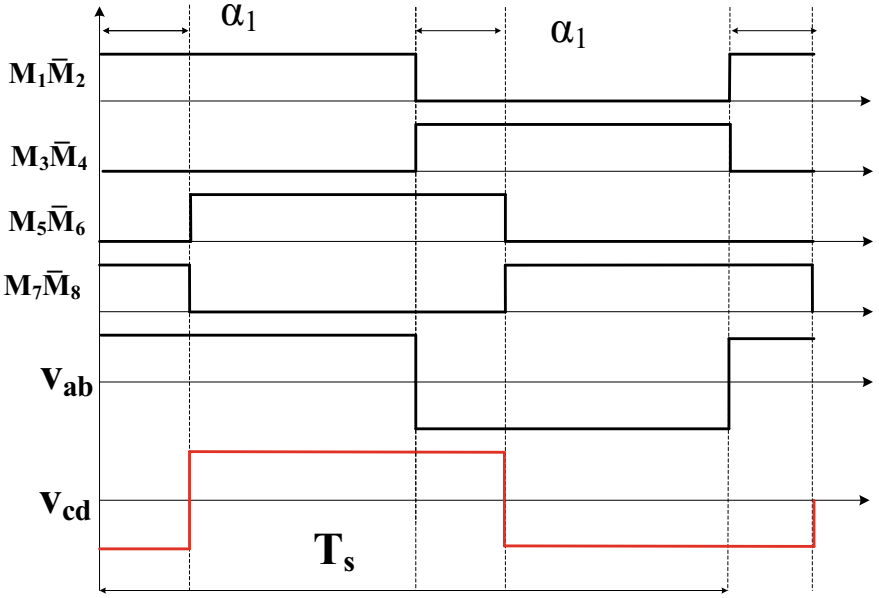


Fig. 4 Waveform of SPS control for DAB converter

4 Switching Scheme

The proposed DAB converter's switching method is shown in Fig. 5. The data acquisition system obtains the output voltage V_2 and output current I_2 in this method. V_2 is compared to $V_2{}^*$ which is reference output voltage and then sent through the PI-controller. The output of this PI is referenced current $I_{\text{ref}}{}^*$, which is compared to I_2 and then passed to another PI-controller. The output of this PI is used for PWM generation for IGBTs of DAB converter.

5 Simulation Parameter and Result

The parameters for constructing DAB converter in MATLAB Simulink are depicted in Table 1.

The results of this model, which was developed in MATLAB Simulink, are listed below. The switching pulses for primary bridge h_1 and secondary bridge h_2 are shown in Figs. 6 and 7, respectively. The phase shift between bridges can be seen to get stable output voltage. The voltage across the primary bridge V_{ab} and secondary bridge V_{cd} and inductor current i_L of HFT is shown in Fig. 8. Both the voltages are square waveforms and phase shifted to each other to transfer the power, and i_L is a triangular waveform due to the voltage being square wave. Figure 9 shows the output

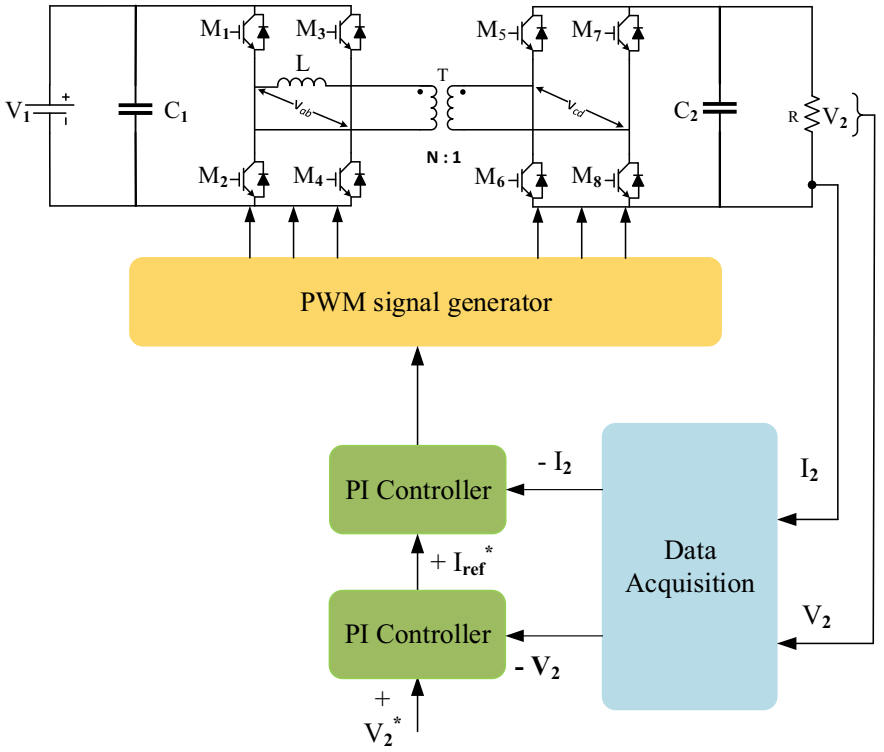


Fig. 5 Switching scheme for DAB converter

Table 1 Simulation parameter

S. No	Parameter name	Value of parameter
1	Input voltage	3800 V
2	Turn ratio	220:30
3	Switching frequency	3000 Hz
4	Total leakage inductance	0.150 mH
5	Capacitor	350 μF
6	Output voltage	200 V
7	Load	65 Ω

voltage of DAB converter by SPS control technique which is stable DC output. The phase shift between bridges is developed from dSPACE 1103 platform as shown in Fig. 10.

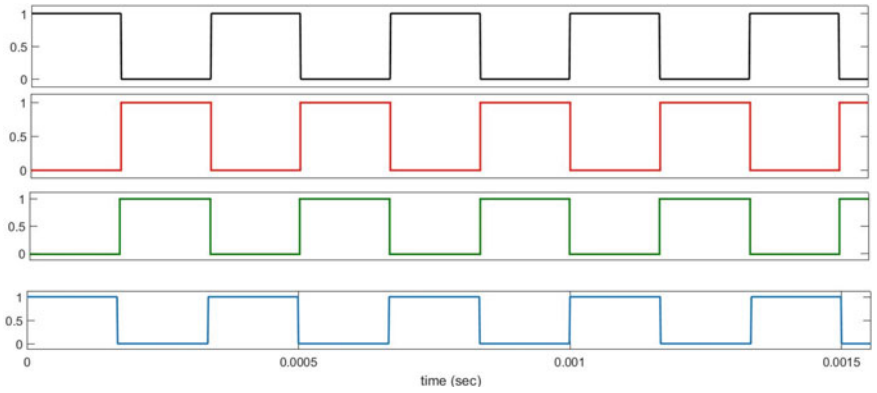


Fig. 6 Pulses for switches for h_1-bridge

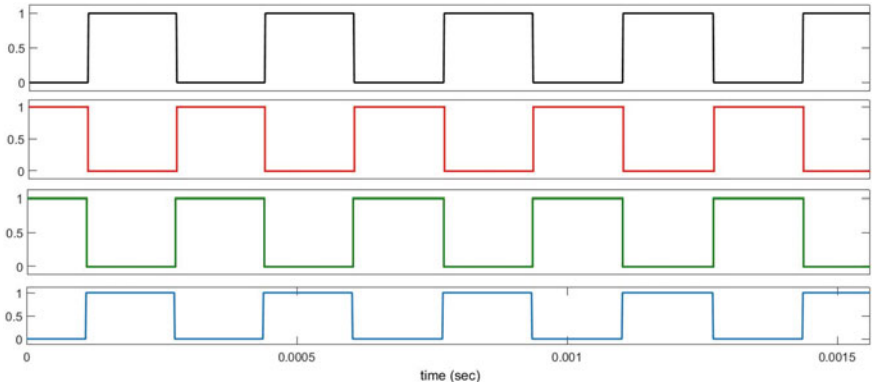


Fig. 7 Pulses for switches for h_2-bridge

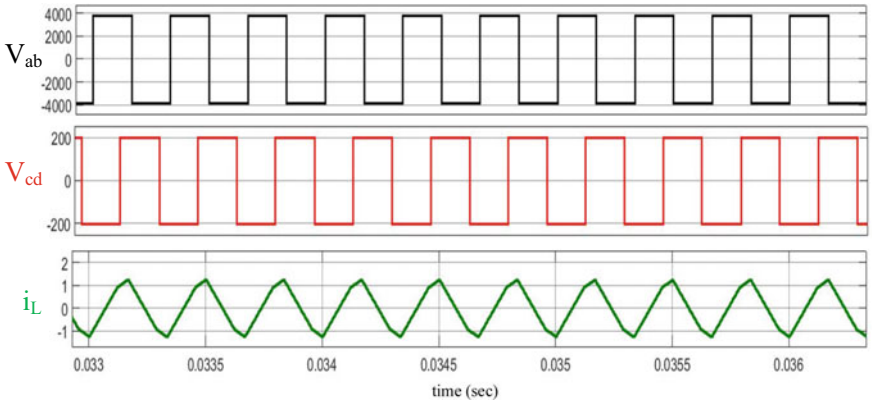


Fig. 8 Waveform of primary and secondary voltage of HFT and inductor current

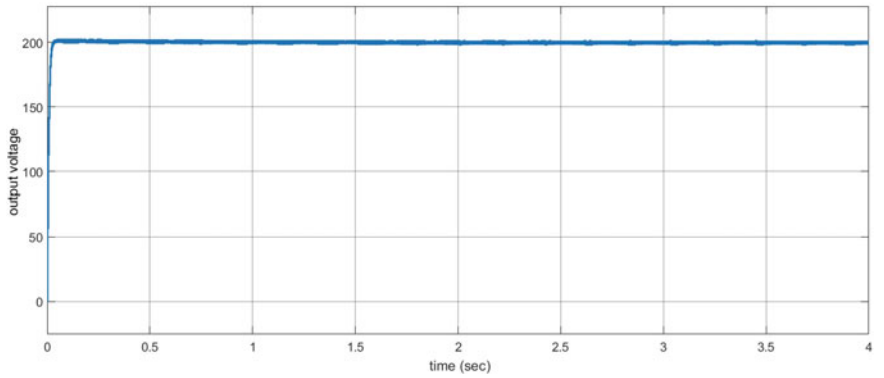


Fig. 9 Output voltage waveform

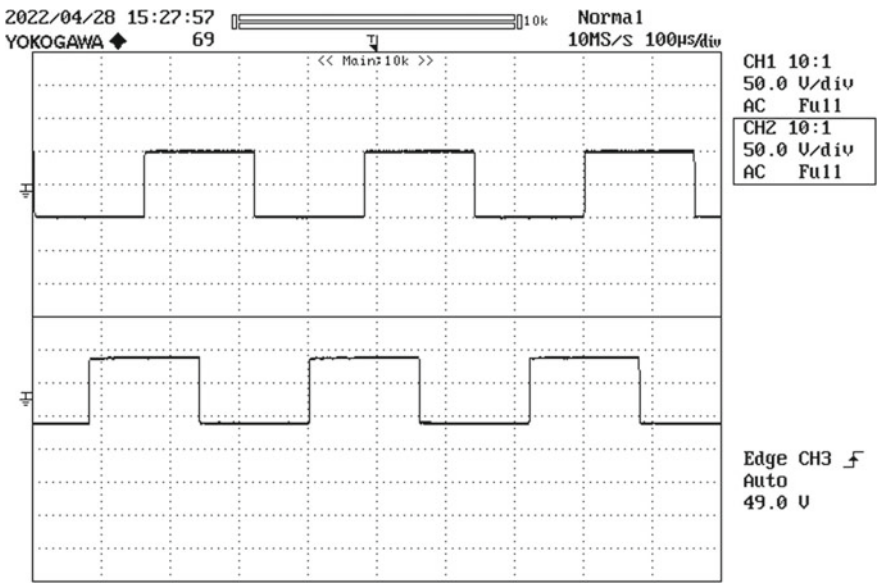


Fig. 10 Phase-shifted pulse from dSPACE 1103

6 Conclusion

SST without control technique has many disadvantages such as high losses and low efficiency. To resolve this problem, DAB converter with SPS control is used to improve the performance of SST. This technique increases the efficiency in whole power range, gives stable DC output with reducing large capacitors, and reduces number of passive components. In this paper, eight IGBT-based DAB converter with HFT is simulated, and phase-shifted pulse is verified from dSPACE 1103 platform.

Two PI controllers are used with SPS control to provide better steady-state performance. It ensures that SPS control is a very simple technique and can be used in high-power applications.

References

1. Zhao T, Yang L, Wang J, Huang AQ (2007) 270 kVA solid state transformer based on 10 kV SiC power devices. In: IEEE electric ship technologies symposium, pp 145–149. https://doi.org/10.1109/ESTS.2007.372077
2. Bignucolo F, Bertoluzzo M, Fontana C (2015) Applications of the solid state transformer concept in the electrical power system. In: 2015 AEIT international annual conference (AEIT), pp 1–6. https://doi.org/10.1109/AEIT.2015.7415235
3. Shadfar J, Pashakolaei HG, Foroud MA, Asghar (2021) Solid-state transformers: an overview of the concept, topology, and its applications in the smart grid. Int Trans Electr Energ Syst 31(9):e12996. ISSN 2050-7038. https://doi.org/10.1002/2050-7038.12996
4. She X, Huang AQ, Burgos R (2013) Review of solid-state transformer technologies and their application in power distribution systems. IEEE J Emerg Sel Topics Power Electron 1(3):186–198. https://doi.org/10.1109/JESTPE.2013.2277917
5. Falcones S, Mao X, Ayyanar R (2010) Topology comparison for solid state transformer implementation. In: IEEE PES general meeting, pp 1–8. https://doi.org/10.1109/PES.2010.5590086
6. Fan H, Li H (2010) High frequency high efficiency bidirectional dc-dc converter module design for 10 kVA solid state transformer. In: 2010 twenty-fifth annual IEEE applied power electronics conference and exposition (APEC), pp 210–215. https://doi.org/10.1109/APEC.2010.5433670
7. Patterson OD, Divan DM (1991) Pseudo-resonant full bridge DC/DC converter. IEEE Trans Electron 6(4):671–678. https://doi.org/10.1109/63.97767
8. Jayaswal K, Palwalia DK (2018) Performance analysis of non-isolated DC-DC buck converter using resonant approach. Eng Technol Appl Sci Res 8(5):3350–3354
9. Zhao B, Song Q, Liu W, Sun Y (2014) Overview of dual-active-bridge isolated bidirectional DC–DC converter for high-frequency-link power-conversion system. IEEE Trans Power Electron 29(8):4091–4106. https://doi.org/10.1109/TPEL.2013.2289913
10. De Doncker RWAA, Divan DM, Kheraluwala MH (1991) A three-phase soft-switched high-power-density DC/DC converter for high-power applications. IEEE Trans Ind Appl 27(1):63–73. https://doi.org/10.1109/28.67533
11. Ríos SJ, Pagano DJ, Lucas KE (2021) Bidirectional power sharing for DC microgrid enabled by dual active bridge DC-DC converter. Energies 14(2):404. https://doi.org/10.3390/en14020404
12. Chen X, Xu G, Han H, Liu D, Sun Y, Su M (2021) Light-load efficiency enhancement of high-frequency dual-active-bridge converter under SPS control. IEEE Trans Ind Electron 68(12):12941–12946. https://doi.org/10.1109/TIE.2020.3044803
13. Meena K, Jayaswal K, Palwalia DK (2020) Analysis of dual active bridge converter for solid state transformer application using single-phase shift control technique. In: 2020 international conference on inventive computation technologies (ICICT), pp 1–6 (2020). https://doi.org/10.1109/ICICT48043.2020.9112398
14. Meena K, Palwalia DK (2020) Voltage control of dual active bridge converter for CO-amorphous core material based solid-state transformer application. IJIREEICE 8(10):33–39. https://doi.org/10.5281/zenodo.4319514
15. Fei X, Feng Z, PuQi N, Xuhui W (2018) Analyzing ZVS soft switching using single phase shift control strategy of dual active bridge isolated DC-DC converters. In: 2018 21st international conference on electrical machines and systems (ICEMS), pp 2378–2381. https://doi.org/10.23919/ICEMS.2018.8549406

16. Jayaswal K, Palwalia DK, Jain G, Kumar P (2015) Design-oriented analysis of non-isolated DC-DC buck converter. Ciencia e Tecnica Vitivinicola journal 30(2):177–213
17. Jayaswal K, Palwalia DK (2018) Performance analysis of non-isolated DC-DC buck converter using resonant approach. Eng Technol Appl Sci Res 8(5):3350–3354
18. Jayaswal K, Palwalia DK (2022) Role of reliability assessment in Si-based non-Isolated DC-DC power electronic converters. Mater Today Proc. ISSN 2214-7853. https://doi.org/10.1016/j.matpr.2022.02.241

Develop a Data Analytics Model for Employee's Performance at Workplace to Increase the Productivity

Rishabh Sinha, Mukunth Narayanan, and Sunil Dhal

Abstract In the present age, job satisfaction of employees is considered very important because it is a major contributor to an organization's productivity and profitability. Many leading and reputed surveys indicate that HR managers find it difficult to identify the various factors of employees' job satisfaction. This paper aims to provide the major factors of employee job satisfaction and also a predictive model that predicts the changes in level of employee work productivity when changes are made to the values of factor(s) of employee job satisfaction. In the first phase (factor identification), 50 factors were obtained from our survey and were subject to factor analysis for dimension reduction. As a result, 50 factors were reduced to 14 major factors or 22 independent variables which were clubbed to 14 major factors. In the second phase, a predictive model is formulated by using the multiple regression technique. The model has 78.8–84.7% accuracy with 95% confidence level. This model helps the HR managers to predict the work productivity of employees within the organization if the value of the factors affecting job satisfaction is increased/decreased. For this part of the project, the newly extracted 22 variables which were clubbed to 14 factors during the factor analysis were taken. In summarization, this study will help HR managers to identify and reduce the employee dissatisfaction thereby aiding with the employee retention and the work productivity of the employees within an organization.

Keywords Job_Satisfaction · Employee_Retention · Work_Productivity · Factor analysis · Predictive model · Multiple regression technique

R. Sinha (✉) · M. Narayanan · S. Dhal
Sri Sri University, Cuttack, India
e-mail: rishabh.s2020mba@srisriuniversity.edu.in

M. Narayanan
e-mail: mukunth.s2021mba@srisriuniversity.edu.in

S. Dhal
e-mail: sunildhal@srisriuniversity.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_62

1 Introduction

The world we know is constantly evolving whether it is from the Stone Age to the modern era or steam engine to electric trains. We humans have always managed to constantly evolve not only ourselves but also the world around us. However, the same can be said for the current organizations or institutions in the world. There was a time when employees were just treated as work slaves who did not have any say in the organization. They can either work or leave; their satisfaction or dissatisfaction with the job wasn't even a consideration for the organization. However, in the current era, job satisfaction of an employee matters a lot as from years of research and proof gathered by various researchers and scholars; it became evident that "The people are the ones who run the organization and not the other way around." Through various studies and research, we now know that organizations, where employees are satisfied with their job, are more profitable than the ones that do not pay much attention to employee satisfaction. All the top companies in the world, like Google, Microsoft, or Dell, go above and beyond just to keep their employees deeply satisfied with them. They provide a lot of benefits that an average company does not and that is also the very reason why people try extra hard to get into such companies. To put it in a nutshell, the job satisfaction of an employee matters a lot for any organization, not only from a profitability point of view but also for enhancing organizational productivity and retention of skilled employees within the organization.

Now, coming to the real-world scenario, various surveys done by major sources like Times Jobs or the Economics Times indicate that HR managers face challenges when it comes to identifying the major factors that impact employee satisfaction within an organization. An HR manager cannot focus on all the parameters that impact the job satisfaction within an organization; hence, our study is to identify the major factors that will help HR managers overcome the day-to-day challenges of the organization. In addition, we are trying to build a model that will predict the productivity of the employees within an organization based on the changes that are made in the values of factors that are obtained through factor analysis in the first half of the project; these factors have a major impact on the job satisfaction within the organization. A simple example to explain this would be, if we increase a single job satisfaction factor, say salary by 30%, and the rest of the factors remain the same how much will the productivity of employees increase or decrease is what our predictive model should predict.

2 Literature Review

Job satisfaction is observed when the reward obtained by an individual for their work matches what they would be expected to achieve [1]. It is found that the performance of employees increases with their level of job satisfaction and showed that job satisfaction is significantly correlated with the occupation of employees as opposed to

gender, qualification, and marital status [2]. However, a study pertaining to Jordanian employees' satisfaction showed that male employees tend to stay in the same job as females, thus citing gender also as a factor. This study also highlighted salary as the most important factor followed by benefits, job security, work stress, and faster promotion. It also cited the factors influencing attrition, wherein low salary featured first followed by bad working conditions and a negative environment [3]. Kathure also affirms that basic pay is considered the most important factor in job satisfaction followed by development opportunities, working conditions, and empowerment [4]. Other factors such as job misfit based on over qualification and the quality of on job training (OJT) were a few of the major factors influencing job satisfaction [5]. A study done to assess employee dissatisfaction in the Bangladesh perspective reported that the main factors causing dissatisfaction was work hours, supervisors, security of income (future), working environment, and administration [6]. Nwobia and Alijohani [7] provide a conceptual model to understand the effect of job dissatisfaction and workplace bullying on employee turnover intention, wherein organizational climate and group cohesion as taken as moderators, citing that these both are also potential factors. A study on the factors influencing job satisfaction of academic staff in Malaysia shows that the quality of top management leadership is also an important factor along with pay and benefits and work environment [8]. A study to assess the impact of an extrinsic factor on job satisfaction at higher education institutions in Vietnam showed that the nature of the job itself is an important factor along with recognition, salary, and supervision [9]. A study conducted to assess the job dissatisfaction of Turkey-based academic and administrative staff showed that lack of transportation service is the most important factor of job dissatisfaction followed by unsatisfactory assignments for both the groups [10].

A study regarding the satisfaction of employees across different business sectors showed that the outsourcing business sector and insurance business sector expressed high levels of dissatisfaction compared to the travel and tourism sector, education sector, etc. [11]. A study conducted to assess the job satisfaction of pharmaceutical companies showed that relationship with co-workers was the most significant factor along with the fairness of work, work condition, and level of pay and promotion. It also reported that the level relationship with the immediate superior had the least impact but still is considered to be a factor [12]. A study conducted for assessing the dissatisfaction in the IT sector revealed that work exhaustion is the main reason for job dissatisfaction [13].

3 Research Objective

This research mainly has two objectives; the first one is as follows: To identify and study the major factors that affect job satisfaction within the organization. The second objective is as follows: To build a predictive model that will examine the degree of increase or decrease in employees' work productivity within an organization with respect to the factors that affect the job satisfaction in the organization.

4 Research Methodology

For the purpose of this research, we will have to design a questionnaire that is based on factors that affect the job satisfaction and productivity of the employees within an organization. We will then conduct a survey using a random sampling technique and go for the reliability test (Cronbach alpha test). After the successful completion of the test with a good score of 0.7 above, we will then proceed toward completing our two main objectives of the research paper.

For the first phase of the research objective, we are going to conduct the factor analysis on all the factors collected and reduce the dimension of the number of factors to just a few major factors based on the responses taken. In the second phase of the objective, we will be building a predictive model to predict, to what degree will increase or decrease happen in the productivity of the employees if one of the factors affecting job satisfaction is changed. This creation of a predictive model will help fill in for the problem that HR managers are facing in the area of enhancing work productivity within the organization.

5 Data Analysis

The data analysis part of this research paper has been divided into 3 segments with the first segment being the reliability test, the second being the factor analysis, and the third being the regression analysis. Now, let us look at all the three segments one by one as follows.

In our research, we have performed the Cronbach's alpha test in order to measure internal consistency. In simple words, we are measuring how closely related a set of items is to a group. Here, in our study, we have used Cronbach's alpha test on the factors that affect the job satisfaction in an organization. There are a total of 51 variables that are taken for this test; all these variables are the factors that are said to have an impact on the job satisfaction in an organization.

Now, coming to the result of the Cronbach's alpha test which is shown in FIg. 1, here, we are getting a value of 0.932, which is above a minimum cutoff value for data reliability of 0.7 in the case of Cronbach's alpha test, and hence, we can say that our data are, highly reliable.

Fig. 1 Cronbach alpha test

Reliability Statistics

Cronbach's Alpha	N of Items
.932	50

5.1 Objective 1: To Identify and Study the Major Factors that Affect Job Satisfaction Within the Organization

Through our study, we have identified 50 factors that impact job satisfaction within an organization. As we know that it is very difficult for an HR manager in an organization to focus on all 50 factors. So, in order to reduce the dimension of a number of factors that affect job satisfaction, we are conducting a factor analysis. Through the use of dimension reduction techniques like factor analysis, we were able to reduce the factors from 50 to just 14 major factors that impact job satisfaction. This reduction is done so that HR managers can focus on these specific 14 factors as they are the most crucial among the lot. From the factor analysis, we also get to know that these 14 factors represent 78.387% of the total variance, which simply means that 78.387% of the total variance can be explained with the help of factors extracted, whereas 21.613% of the total variance still remains unexplained. The following are various tables or matrices obtained through factor analysis on the 50 factors that impact job satisfaction in an organization. Figure 2 given indicates that there are a total of 50 factors or parameters that impact job satisfaction within an organization. These parameters are on the Likert scale. The value of Likert scale values is present between one and five. In addition, the result of factor analysis shows that there were about 14 factors extracted from the total of 50 factors and that it represents 78.387% of the total variance. This means that 78.387% of the total variance is explained through the help of these 14 factors.

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	12.938	25.876	25.876	12.938	25.876	25.876	5.012	10.024	10.024
2	3.517	7.034	32.910	3.517	7.034	32.910	4.313	8.627	18.650
3	3.263	6.525	39.436	3.263	6.525	39.436	3.908	7.817	26.467
4	3.012	6.023	45.459	3.012	6.023	45.459	3.485	6.969	33.437
5	2.643	5.287	50.745	2.643	5.287	50.745	3.282	6.563	40.000
6	2.137	4.274	55.020	2.137	4.274	55.020	2.768	5.536	45.535
7	1.907	3.813	58.833	1.907	3.813	58.833	2.620	5.241	50.776
8	1.711	3.422	62.255	1.711	3.422	62.255	2.562	5.124	55.900
9	1.627	3.254	65.509	1.627	3.254	65.509	2.143	4.286	60.186
10	1.444	2.889	68.398	1.444	2.889	68.398	2.063	4.126	64.311
11	1.372	2.744	71.142	1.372	2.744	71.142	2.040	4.079	68.391
12	1.304	2.608	73.750	1.304	2.608	73.750	1.858	3.715	72.106
13	1.197	2.393	76.143	1.197	2.393	76.143	1.822	3.644	75.750
14	1.122	2.243	78.387	1.122	2.243	78.387	1.318	2.637	78.387
15	.965	1.930	80.317						
16	.928	1.856	82.172						
17	.865	1.729	83.902						
18	.778	1.556	85.458						
19	.727	1.454	86.912						
20	.682	1.364	88.275						
21	.605	1.210	89.485						
22	.547	1.094	90.579						
23	.500	.999	91.578						
24	.441	.882	92.460						
25	.410	.820	93.280						
26	.378	.757	94.036						

Fig. 2 Total variance explained

Figure 3 given indicates the Pearson correlations between the items and components. It is evident from the table that 14 parameters hold significant influence and impact on job satisfaction within an organization, and also, these factors can be defined through the rotation matrix given below:

5.1.1 Extracted Factor Table

See Fig. 4.

5.1.2 Objective (1) Findings

As seen in Fig. 4, we have been able to extract 14 factors from the 50 factors that affect job satisfaction within the organization. Now, coming to the process of extracting the factors, first, we look at the rotated component matrix. We start by looking at column 1 which indicates factor number 1. Then, we check if there are multiple factors loading values equal to or above 0.7. If there are multiple values above 0.7 under the same column, then we combine the factors associated with those factors loading values. In our case, this is done for factors numbers 1, 2, 3, 4, 5, 7, 8, and 9. Furthermore, if there are no multiple factors loading values equal to or above 0.7 in the entire column, then the factor against the highest factor loading value in the column will be the only factor for that column. In our case, factors number 6, 10, 11, 12, 13, and 14 were the ones that have only one highest value above 0.7 or only one highest value in the entire column. So, this is how we have extracted our 14 factors out of all 50 factors that affect job satisfaction within organizations.

5.2 Objective 2: To Build a Predictive Model that will Examine the Degree of Increase or Decrease in Employees' Work Productivity within an Organization with Respect to the Factors that Affect the Job Satisfaction in the Organization

To analyze the change and the level of influence of productivity on the factors affecting job satisfaction, we have built a predictive model. This predictive model like all the other models also has two types of variables, dependent and independent variables. In our case, the dependent variable is "Employee Work Productivity," and the independent variable are all the 50 variables that affect the job satisfaction within the organization. This can be depicted by the predictive model/table given below.

The model summary table given below indicates, how much of a good fit is our predictive model. We check for the R-squared and adjusted R-Squared values to see

Rotated Component Matrix[a]

	Component													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Service	-0.007	0.107	0.648	-0.024	0.016	0.195	0.157	0.161	-0.077	-0.377	-0.044	-0.255	0.072	-0.167
Salary	0.217	0.228	-0.026	-0.090	0.002	-0.125	0.787	-0.061	0.149	-0.084	-0.073	0.041	0.009	0.050
Effective appraisal system	0.077	0.190	0.158	-0.042	0.095	-0.117	0.074	0.743	0.095	-0.045	0.146	0.104	-0.197	0.073
Merit based promotion	-0.025	0.085	0.113	0.246	0.079	0.009	0.025	0.723	0.006	0.270	-0.012	-0.084	0.034	-0.120
Incentives	-0.029	-0.046	0.001	0.059	0.184	0.147	0.818	0.153	-0.137	0.120	0.017	-0.183	-0.006	0.010
Proper training	0.381	0.074	0.188	0.162	0.360	0.021	0.504	0.060	0.845	0.253	0.138	0.030	0.100	0.007
Career growth opportunities	-0.122	0.062	-0.007	0.117	0.168	0.165	0.091	0.139	-0.091	0.845	0.087	-0.037	0.019	-0.013
Quality of management	0.382	0.017	0.284	0.246	0.194	-0.036	0.210	0.342	-0.166	0.325	0.169	-0.303	-0.152	0.172
Worklife balance	-0.016	0.214	0.070	0.237	0.689	-0.195	0.069	0.198	-0.156	0.156	-0.195	0.199	-0.069	-0.148
Work culture	0.428	0.091	0.016	0.189	0.635	-0.255	-0.063	0.019	0.070	0.240	0.021	0.089	0.112	0.008
Quality of reporting manager	0.072	-0.098	0.002	0.188	0.096	0.144	-0.108	0.038	-0.177	0.023	0.782	0.286	0.150	-0.038
Adequate resources	-0.041	0.024	0.283	0.144	0.820	0.057	0.199	0.114	0.062	-0.031	0.130	0.049	-0.115	-0.079
Opportunities for learning new Skills	0.077	0.046	0.150	-0.127	0.739	0.086	0.172	-0.022	-0.014	0.028	0.307	-0.154	0.108	0.279
Quality of work assigned	0.115	0.092	0.279	0.123	0.198	-0.049	0.346	0.268	0.073	0.217	0.456	-0.346	-0.057	0.032
Rewards and Recognition	0.081	0.302	0.131	0.042	0.180	0.083	0.182	0.430	0.272	0.105	0.489	0.071	-0.089	0.179
Flexibility of work hours	0.145	0.369	0.457	0.075	0.313	0.135	-0.014	0.068	0.161	0.215	0.117	0.202	-0.001	-0.075
Gender discrimination	0.182	0.082	0.705	-0.195	0.220	0.166	0.004	0.184	-0.171	-0.035	0.146	0.093	-0.064	0.110
Logistics	0.287	0.026	0.294	-0.014	0.165	0.338	0.031	0.401	-0.031	0.195	-0.292	0.175	0.103	0.310
Job Role	0.268	0.060	0.212	0.061	0.044	0.128	0.290	0.116	0.265	0.359	-0.009	-0.315	0.440	-0.185
Quality of communication	0.408	-0.085	0.549	0.273	0.128	0.175	0.130	0.183	0.176	0.112	0.159	-0.096	-0.006	-0.029
Influx of technology	0.129	-0.129	-0.070	0.136	-0.131	0.182	0.028	0.105	0.778	-0.182	-0.101	-0.095	0.138	0.113
Job security	0.003	0.327	0.044	-0.039	0.108	0.012	0.047	0.001	0.774	0.088	0.031	-0.069	-0.166	-0.031
Biased work environment	0.302	0.376	0.452	0.341	0.234	-0.134	-0.182	0.015	-0.085	0.054	0.103	-0.116	-0.331	0.020
Work stress	0.114	0.191	0.029	0.025	0.107	0.075	-0.112	0.068	-0.143	-0.025	0.188	0.808	-0.053	0.033
Quality of job performance	0.040	0.698	0.108	0.273	0.117	0.114	0.172	0.153	0.004	-0.176	0.149	0.177	-0.166	-0.320
Meaning and purpose in job	0.060	0.225	0.083	0.171	-0.021	0.845	-0.030	-0.053	0.077	0.080	0.050	-0.022	0.071	-0.001
Sense of identity	0.314	0.373	-0.049	0.035	0.308	0.385	-0.171	0.030	-0.159	-0.044	-0.099	0.049	-0.427	-0.314
Job Location	0.629	-0.147	0.249	0.074	0.184	0.343	-0.007	-0.033	0.108	0.089	0.196	-0.017	0.019	-0.099
Organizational relationship	-0.023	0.732	0.123	-0.117	-0.012	-0.024	0.051	0.007	0.200	-0.093	-0.090	-0.233	0.266	0.124
Disengagement within the organization	0.373	0.469	0.251	0.235	0.338	0.152	-0.198	0.190	-0.061	0.050	0.024	-0.199	-0.044	0.102
Job passion	0.110	0.754	0.133	0.170	0.032	0.151	0.117	0.199	0.056	0.231	0.021	0.171	0.075	0.093
Work Ethics	0.549	0.467	0.055	0.080	-0.149	0.187	0.206	0.299	0.079	0.072	0.147	0.190	0.078	-0.108
Work Motivation	0.352	0.711	-0.042	0.105	0.158	0.113	0.071	0.053	-0.023	0.100	-0.062	0.186	-0.200	0.093
Devaluation of an employee	0.097	0.178	0.048	0.574	0.203	0.135	-0.059	0.218	0.109	0.166	0.407	-0.005	-0.207	-0.076
Excessive Authority	0.161	0.036	0.186	0.726	0.092	0.078	0.075	-0.064	0.016	0.210	0.046	0.011	0.142	0.097
Ideas and Innovation	0.129	0.328	-0.176	0.729	0.001	0.315	0.142	0.174	-0.017	-0.116	0.050	0.032	0.004	0.088
Unchanged Daily Work Routine	0.308	-0.136	0.240	0.520	0.252	0.261	-0.192	0.204	0.241	-0.004	0.027	-0.052	0.134	-0.111
Mix up of professional and private life	0.256	0.078	0.100	0.154	0.028	0.246	-0.047	-0.175	-0.111	-0.023	0.082	-0.009	0.793	0.022
Transparency of the information	0.542	0.286	0.067	0.439	0.192	0.223	-0.179	0.274	0.066	0.176	0.067	0.098	0.075	-0.236
Employee Harassments	0.360	0.251	0.174	0.213	-0.008	0.204	-0.170	0.137	0.097	0.353	0.095	0.106	-0.276	0.408
Collecting highly sensitive information about their customers	0.116	-0.182	0.358	0.295	0.157	0.363	-0.006	-0.244	0.224	0.116	0.229	0.064	0.057	0.118
Job is creating environmental and health problems	0.513	0.271	0.321	0.269	0.070	0.156	-0.087	-0.108	-0.255	0.030	0.303	-0.198	0.045	0.063
Workplace rigidity	0.098	0.434	0.509	0.449	0.063	-0.018	-0.038	-0.079	-0.264	0.002	0.128	0.175	0.073	0.164
Retirement plan	0.874	0.217	0.184	0.000	0.017	0.054	0.149	0.041	0.008	-0.076	-0.036	0.116	-0.003	0.133
Organization policy	0.780	0.264	0.394	0.403	0.045	0.163	-0.004	-0.003	0.066	-0.033	0.179	0.042	0.046	-0.088
Office Ambience	0.604	0.074	-0.048	0.131	0.018	-0.012	0.250	-0.037	0.066	-0.181	-0.098	0.004	0.217	0.184
Work from home	0.336	0.004	0.419	0.222	0.012	-0.115	0.271	-0.308	0.023	0.204	-0.063	0.374	-0.019	-0.106
Work Place productivity among the colleagues	0.152	0.141	0.718	0.211	0.140	0.024	-0.059	0.132	0.157	0.167	-0.143	0.012	0.237	0.131
High Skillset development	0.250	0.268	0.099	0.216	0.097	0.333	0.363	0.007	0.165	-0.077	0.020	0.029	0.078	0.505
Overall Organizational Performance	0.109	0.093	0.213	0.194	-0.234	0.887	0.138	0.027	0.121	0.169	0.150	0.117	0.216	0.159

Fig. 3 Rotated component matrix

Factor Number	Factor Name	Eigen Value	% of Variance	Statements	Loading
1	Organizational policy with Retirement benifits inclusion	0.031	0.062	Retirement plan	0.874
		0.28	0.056	Organization policy	0.78
2	Job passion & Motivaton	0.232	0.465	Job passion	0.754
		0.19	0.379	Work Motivation	0.711
3	Organizational Productivity and Decipline	0.865	1.729	Gender discrimination	0.705
		0.008	0.015	Work Place productivity among the colleagues	0.718
4	Ideas, Innovation and Power impartiality	0.141	0.282	Ideas and Innovation	0.729
		0.16	0.32	Excessive Authority	0.726
5	Organizational Resources	1.304	2.608	Adequate resources	0.82
		1.197	2.393	Opportunities for resources for learning new skills	0.739
6	Value in Job	0.378	0.757	Meaning and purpose in job	0.845
7	Salary & Incentives	3.517	7.034	Salary	0.787
		2.643	5.287	Incentives	0.818
8	Effective Employee Performance Measurement System	3.263	6.525	Effective appraisal system	0.743
		3.012	6.023	Merit based promotion	0.723
9	Technology Rise & Job Security	0.605	1.21	Influx of technology	0.778
		0.547	1.094	Job security	0.774
10	Career Growth Opportunities	1.907	3.813	Career growth opportunities	0.845
11	Quality of Reporting Manager	1.372	2.744	Quality of reporting manager	0.782
12	Work Stress	0.441	0.882	Work stress	0.808
13	Mix-Up of Professional and Private Life	0.113	0.226	mix up of professional and private life	0.793
14	Advanced Skill Development	0.006	0.012	High Skillset development	0.505

Fig. 4 Factor extraction

if our model is a good fit or not. For a model to be a good fit, it should be adjusted R-squared can also be called a modified version of R-squared. The value of a good fit model for R-squared and adjusted R-squared should be at least 0.7 or above. This means that the higher the value of R-squared and adjusted R-squared the better the predictive model would be.

In Fig. 5, the R-squared and adjusted R-squared values are 0.847 and 0.788, which can also be interpreted as that our model accuracy is about 78.8–84.7%, which indicates that our predictive model is a good fit.

The ANOVA table shows whether our model is significant or not. For a regression model to be significant, its P-value should be less than 0.05. The *P*-value of less than 0.05 shows a 95% confidence that the slope of the regression line is not zero. This can be interpreted as the two variables that are dependent and independent variables having a significant linear relationship with each other.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.920[a]	.847	.788	7.141

Fig. 5 Accuracy of the model determined by R-square and adjusted R-square values

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	36628.266	50	732.565	14.366	<.001[b]
	Residual	6629.005	130	50.992		
	Total	43257.271	180			

Fig. 6 ANOVA table

In Fig. 6 given above, the *P*-value is < 0.001 which indicates that our dependent and independent variables have a significant linear relationship with each other. The coefficient table provides the coefficient values of all the independent variables along with their respective *P*-value, standardize coefficient beta value, *T*-value, standard error, and unstandardized *B* value. This table provides us with the coefficient values that will be useful for our predictive model. The coefficient values for building our "Workplace Productivity" predictive model have been directly taken from the coefficient table obtained in the SPSS software and are directly serving as the coefficient values for the independent variable in the "Workplace Productivity" predictive model.

The predictive model that we are building in our case has a total of 50 independent variables; however, we are building the model based on the 14 major factors or variables that we extracted from those 50 variables during the factor analysis. Thou, the 14 factors have been obtained due to clubbing during the factor analysis; the actual number of independent variables comes out to be 22. Hence, these 22 major independent variables are taken for building the predictive model, and the reason for the same is that they are the factors that majorly affect the entire work productivity within organization. In addition, focusing on all the 50 independent variables is a very difficult job for any HR manager. Hence, we have reduced the independent variable from 50 to 22 for building our predictive model by the use of factor analysis which is taken as a part of objective 1. The predictive model can be mathematically represented as follows:

Work Productivity = 83.854 + Salary * (−1.236) + Effective Appraisal System * (16.648) + Merit Based Promotion * (−24.072) + Incentives * (−21.702) + Career growth * (−1.385) + Quality of Reporting Manager * (−21.013) + Adequate Resources * (−4.905) + Opportunities for Learning new skill * (−17.829) + Gender Discrimination * (−4.207) + Influx of Technology * (−11.794) + Job Security *

(12.250) + Work Stress * (5.566) + Meaning and Purpose in job * (−3.135) + Job Passion * (−10.933) + Work Motivation * (15.222) + Excessive Authority * (11.819) + Ideas and Innovation * (−7.500) + Mix up of professional and Private life * (9.092) + Retirement Plan * (−23.163) + Organizational Policy * (−11.197) + Work Place productivity among colleagues * (1.607) + High Skillset Development * (−4.039). The abovementioned mathematical expression is our predictive model to find out the increase or decrease or the influence on the work productivity of the employee in an organization with respect to the variables that impact or influence the job satisfaction within the organization and in turn affect the work productivity of the employees within organization.

5.2.1 Objective (2) Findings

The predictive model built is able to predict the work productivity of employees within the organization with the help of factors that impact the job satisfaction within an organization. This model will help you find out to what degree of increase/decrease in the factors affecting job satisfaction will have on the work productivity of employees within the organization Now, suppose if you want to find out how much will the work productivity increases/decreases based on the factors affecting job satisfaction, we can study this with an analogy.

Let's say that employees working in an organization have high work stress. So, in order to reduce the work stress, the organization decided to reduce the work stress by 30% in an organization. Initially, the employee on average was working on the 2 projects each day, and a point to be noted is that the employees did not handle the project alone or did the maximum work. The entire work was shared by the employees of the organization. However, even after sharing the workload, it was still high for every person in the organization. So, now coming to the main point, if the HR manager of the organization substitutes the value of the new 30% decreased work stress and the other 49 factors that affect the job satisfaction, then we might find out how will the work productivity within an organization increase/decrease. The same can be done with the other 21 parameters, which means we can increase or decrease all the 22 parameters simultaneously or one by one to check the influence on the work productivity of employees within the organization.

6 Conclusion

The objective of the study is to increase the work productivity within the organization by analyzing the factors that affect the job satisfaction within the organization. In order to improve the work productivity within an organization, we first did an exploratory study on the factors that impact the job satisfaction within an organization. We managed to pin point 50 factors which are responsible for job dissatisfaction

within an organization. However, to observe and consider all the 50 factors is something very difficult for any human being. So, to make things easier for people like HR or HR managers all over the globe, we used the dimension reduction technique called factor analysis. The aid that we received by subjecting all 50 factors that impact the job dissatisfaction within the organization was that we were able to condense all the 50 factors into 14 major factors after clubbing the 22 major independent variables that impact the job dissatisfaction within an organization. Moreover, by using these 22 major variables and multiple regression technique, we were to build our work productivity prediction model. The model basically predicts the change in the work productivity within the organization with respect to the increase/decrease in the factors that impact job satisfaction. Furthermore, through this research, we are able to increase the work productivity within organization by simultaneously decreasing the job dissatisfaction within the organization.

References

1. Magdalena-Camella B, Marius B (2011) Stress management and work performance—a case study. J Eng Stud Res 17(3)
2. Inayat W, Khan MJ (2021) A study of job satisfaction and its effect on the performance of employees working in private sector organizations, Peshawar. Educ Res Int 2021:9. Article ID 1751495, https://doi.org/10.1155/2021/1751495
3. Abuhashesh M, Aldmour R, Masa'deh R (2019) Factors that affect employees' job satisfaction and performance to increase customers' satisfactions
4. Kathure MB (2018) Factors affecting job satisfaction among employees in telecommunication companies in Kenya: a case study of horizon contact center. Published in core.ac.uk
5. Awino M (2017) Job dissatisfaction: an analysis of skill and knowledge mismatch of on-job trained employees in institutions of higher learning. Glob J Human Resour Manag 5(1):91–101. Published by European Centre for Research Training and Development UK
6. Alam J (2015) An empirical study about why dissatisfaction arises among the employees and what it consequences: Bangladesh perspective. Manag Stud Econ Syst 1:171–180. https://doi.org/10.12816/0019379
7. Nwobia I, Aljohani M (2017) The effect of job dissatisfaction and workplace bullying on turnover intention: organization climate and group cohesion as moderators. Int J Mark Stud 9:136. https://doi.org/10.5539/ijms.v9n3p136
8. Hee OC et al (2020) Factors influencing job satisfaction among academic staffs. Int J Eval Res Edu (IJERE) 9(2):285–291 ISSN 2252-8822, https://doi.org/10.11591/ijere.v9i2.20509
9. Nguyen PND, Nguyen LLK, Le DNT (2021) The impact of extrinsic work factors on job satisfaction and organizational commitment at higher education institutions in Vietnam. J Asian Finan Econ Bus 8(8). ISSN 0259-0270
10. Eroğlu ŞG et al (2014) A research on job dissatisfaction of the university staff. Eur J Res Educ 2(Special Issue):185–192. ISSN 2147-6284, https://doi.org/10.15527/ejre.201426574. Published by International Association of Social Science Research—IASSR
11. Sharma SC, Gupta R (2020) Job satisfaction: difference in levels among selected industries. Int J Rec Technol Eng (IJRTE) 8(6). ISSN 2277-3878 (Online)
12. Parvin MM, Kabir MNN (2011) Factors affecting employee job satisfaction of pharmaceutical sector. Aust J Bus Manage Res 1(9):113–123
13. Ramos R, Moro S, Rita P (2020) What drives job satisfaction in IT companies? Int J Product Perform Manag 70:391–407. https://doi.org/10.1108/IJPPM-03-2019-0124

Estimating Related Words Computationally Using Language Model from the Mahabharata an Indian Epic

Vrunda Gadesha, Keyur Joshi, and Shefali Naik

Abstract 'Mahabharata' is the most popular among many Indian pieces of literature referred to in many domains for completely different purposes. This text itself is having various dimension and aspects which is useful for the human being in their personal life and professional life. This Indian 'epic is originally written in the Sanskrit Language. Now in the era of natural language processing, artificial intelligence, machine learning, and human-computer interaction, this text can be processed according to the domain requirement. It is interesting to process this text and get useful insights from Mahabharata. The limitation of the humans while analyzing Mahabharata is that they always have a sentiment aspect toward the story narrated by the author. Apart from that, the human cannot memorize statistical or computational details, like which two words are frequently coming in one sentence? What is the average length of the sentences across the whole literature? Which word is the most popular word across the text, what are the lemmas of the words used across the sentences? Thus, in this paper, we propose an NLP pipeline to get some statistical and computational insights along with the most relevant word searching method from the largest epic 'Mahabharata'. We stacked the different text processing approaches to articulate the best results which can be further used in the various domains where Mahabharata needs to be referred.

Keywords NLP pipeline · Text processing · Word relation in Mahabharata · Word embedding · Word2Vec

V. Gadesha (✉) · K. Joshi · S. Naik
School of Engineering and Applied Science, Ahmedabad University, Ahmedabad, Gujarat, India
e-mail: vrunda.g@ahduni.edu.in

K. Joshi
e-mail: keyur.joshi@ahduni.edu.in

S. Naik
e-mail: shefali.naik@ahduni.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_63

1 Introduction

Natural Language processing is the cutting-edge technology equipped with efficient tools and techniques to deal with unstructured text data. Using NLP pipeline techniques, a large amount of text can be processed very quickly and accurately. The most important point of processing the fictional text using NLP is that the text will be analyzed without adding any sentiments to it. 'Mahabharata' is the story orally often narrated and recreated across the world in different forms. Thus, humans have sentiments attached to them by default. So, to get the computational details about Mahabharata, we used the elements of the NLP pipeline to answer the following questions which do not have any sentiment aspect attached with it.
(1) How rich is 'Mahabharata' in terms of words?
(2) Does the sentence length of 'Mahabharata' distribute normally across the whole literature?
Apart from this, we are addressing the problem that how can we find the *most related* word from such a large text without reading it. In this paper, we are approaching the NLP pipeline followed by the language model which is searching for the most related words from the large text.

2 Literature Review

Mahabharata is a treasure of life lessons. The literature is narrating the phenomena and story which is lived by more than 200 people [1] which has been redacted between 400BCE and 400CE [2]. We can see the glimpse of various events that occurred in past across India and even across the globe [3].

Among these chunks, the city Bishnupur in West Bengal, India is famous for its terracotta temples. These temple's walls are carved with terracotta panels describing various events from 'Mahabharata'. These images are captured and used as a 3D image dataset known as BHID (Bishnupur Heritage Image Dataset) for various computer vision applications [4].

The story of Mahabharata is retold in various art forms like plays, short stories, paintings, poems such as 'Kiratarjunyam' to make people understand the right ethics that not to make difference between 'high-man' and 'low-man' where 'Lord Shiva' himself described as 'Kirat' [5] and translated books in various Indian languages. Though the orality affects the translation, according to paper [6], the translation of literature may be treated as an independent text because 'a study of translation is a study of language', and Mahabharata is retold in various Indian languages. Between these all forms of art, a unique art called 'Wayang (leather puppets)' is famous for recreating the Mahabharata story in Bali - Indonesia. Sudiatmika (et al. 2021) and fellow researchers have classified the 'Mahabharata events' presented in this art form. They used the R-CNN algorithm to achieve the recognition of events and the characters such as 'Wayang Arjuna' and 'Wayang Yudhistira' [3].

This epic is even inspiring for the technologist to create various taxonomies for the fictional domain (TiFI) [7] and launch 'ENTYFI'—the first technique for typing entities in fictional text. This 5-steps technique is useful to generate supervise fiction typing, supervised real-world typing, and unsupervised typing [8]. A large number of events and characters in the epic is also useful for ontology (a knowledge representation structure). In the current scenario, the Web resources are more explored for ontology enrichment rather than the question-answer-pair (QA-pair). Authors in paper [9] applied such QA-pair on the 'Mahabharata domain' to convert them into potential triples (subject, predicate, and object) and identify the triples which are new, more precise, and related to the domain for ontology enrichment in literature.

While looking at the various aspect of Mahabharata, excluding the psychological aspects is not possible. Authors in paper [10] have explored the evidence for the most fundamental metaphor used for the mind - 'The Mind is a Container' in Indian Epic 'Mahabharata' and 'Ramayana' plus the Greek Epic poem 'Homer and Hesiod' to traverse the cognitive phenomena in the epic literature.

Mahabharata Wiki Article is featured in the 100 most viewed Wikipedia article list. It is easy to give the context of the literature to people who belong to different domains. Thus, it is important to have a computational, analytical, and sentimental analysis of the text to get meaningful insights [11]. In [12], the authors have derived interesting insights from the English translation of Mahabharata [1] by applying preprocessing, POS tagging, co-occurrence analysis, sentiment analysis of text and characters, and emotional analysis. The insights which are given about the character and phenomena are versatile enough to use in different domains.

Apart from the technology development, the treatise has relevance to the modern society and is helpful to derive management lessons such as strategic management, creation and relation with powerful friends and Allies, effective leadership style, successful team building, shared goal and ownership of the goal, commitment to the goal, role clarity, understanding the ground realities, and empowering women [13]. The most important part of this epic is 'Bhagavad Gita' said during Bhishma Parva also gives lessons of intrapersonal skills like self-development, sublimation/management of the physical dimensions, sublimation/management of the psychological dimensions, deontology, desire management, anger management, mind management, emotional stability, fear management, self-motivation, empathy, and social welfare [14]. This epic gives the zoom version of the art of concentration with the lifespan of Arjuna. Different events can lead us to derive the factors which can be considered for the concentration like enthusiasm, dedication, aptitude, emotional or physical state, and environment [15]. The epic context is shaping the thinking of society over the centuries. And this is reflected in our modern literature for children and adults. The stories derived from the epic show the disability as a curse or sin, but the modern literature shows the positivity and power of the disability. It portrays the usefulness of disabled people to society. In the context of Mahabharata, the approach toward the disability may fall under the bucket of 'don'ts'[16].

3 Methodology

This paper aims to carve the non-semantic, statistical, and computational insights along with finding the most relevant words from the largest Indian epic 'Mahabharata'. Figure 1 shows the NLP pipeline, which is defined to get robust results on the text. During this experiment, 'The Mahabharata of Krishna Dwaipayana Vyasa—The English translation (1886–1889) by Kesri Mohan Ganguli' is used in '.EPUB' format as a dataset.

3.1 .EPUB File Conversation into Data Structure

The '.EPUB'—(electronic publication) format is a very popular format of the e-book in digital documentation. This format is not only useful to read e-books using multiple devices such as android/mac mobiles, tablets, laptops, or desktops, but these files are also useful for text processing. The EPUB format is released as an archive file built on the XHTML method. The tag format of XHTML can be flattened into any data structure which is readable by machine language. Here, the whole e-book is converted into a Python list data structure.

As shown in Fig. 2, the 'Mahabharata' e-book is divided into sequential data structure. While conversation, the 'new line' is converted into '\ n', and page break

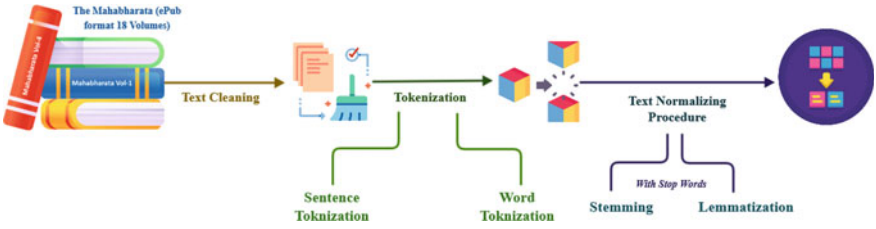


Fig. 1 Text preprocessing pipeline on the Mahabharata

```
In [11]: out = epub2text('mahabharata.epub')

In [12]: out

Out[12]: ['\n \n \n \n \n ',
          'Adi Parva \n \n \n Translators Preface \n Section 1 \n Section 2 \n P
         aushya Parva \n Pauloma Parva \n Astika Parva \n Adivansavatarana Parva
         \n Sambhava Parva \n Jatugriha Parva \n Hidimva-Vadha Parva \n Vaka-Vad
         ha Parva \n Chaitraratha Parva \n Swayamvara Parva \n Vaivahika Parva \
         n Viduragamana Parva \n Rajya-Labha Parva \n Arjuna-Vanavasa Parva \n S
         ubhadra-Harana Parva \n Haranaharana Parva \n Khandava-Daha Parva \n \n
         \n ',
          'Translators Preface \n The object of a translator should ever be to h
```

Fig. 2 Mahabharata—in a list structure

is converted into '\ o'. Apart from this, we do have some unwanted elements such as comma (,), semicolon (;), and apostrophe 's' ('s).

3.2 Text Cleaning

The Mahabharata story contains many punctuations which are important to understand the sentiments for humans, but it not useful for the machine. Text cleaning is addressing the problem to handle unwanted elements. Using Python library 're' (regular expression) and 'string', the redundant elements such as comma (,), semicolon (;), and apostrophe 's' ('s) are removed from the whole text, and the text is now stored a unit string. In the general case, full stop (.) is also removed during the text cleaning of the dataset but this process required full stop (.)

```
In [21]: mahabharata_str_cleaned

Out[21]: '          adi parva       translators preface   section 1   section 2
         paushya parva   pauloma parva   astika parva   adivansavatarana parva
         sambhava parva   jatugriha parva   hidimva-vadha parva   vaka-vadha par
         va   chaitraratha parva   swayamvara parva   vaivahika parva   viduraga
         mana parva   rajya-labha parva   arjuna-vanavasa parva   subhadra-haran
         a parva   haranaharana parva   khandava-daha parva       translators p
         reface   the object of a translator should ever be to hold the mirror u
         pto his author. that being so  his chief duty is to represent so far as
         practicable the manner in which his author s ideas have been expressed
```

Fig. 3 Result of text cleaning process on Mahabharata in unit string format

while performing the next step of the pipeline called tokenization. The reason behind keeping the full stop is to define the end of the sentences. After tokenization, we can find the number of words occupied in each sentence which can be identified as a word distribution pattern (Fig. 3).

3.3 Tokenization

The concept of dividing the text document into small snippets is known as tokenization. The tokenization can be applied in two different ways on the text document: (a) sentence tokenization and (b) word tokenization. These can generate a bunch of sentences, words, phrases, tokens, or symbols [17]. Usually, tokenization is applied as a primary and conventional text preprocessing step in an NLP pipeline.

In the text preprocessing of the Mahabharata, we used the 'natural language toolkit - sent_tokenize()' method to divide the whole text into sentences. The whole Mahabharata is divided into 130,700 sentences with variable lengths. The length distribution is described in Fig. 4.

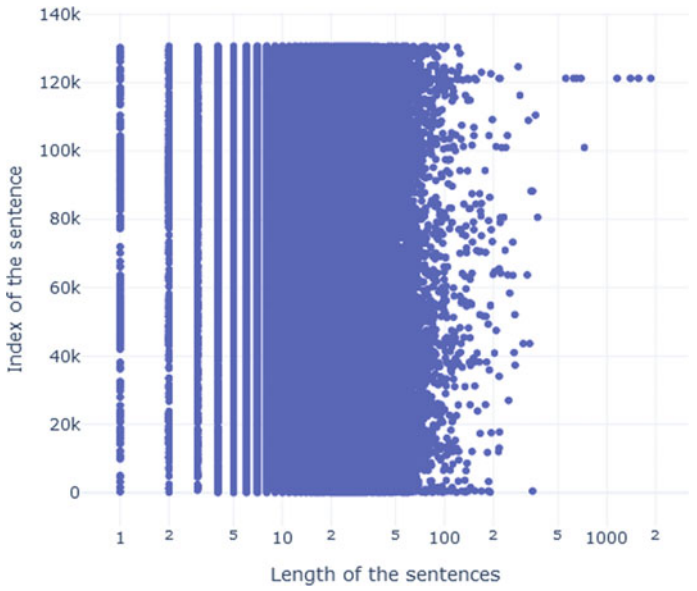


Fig. 4 Sentence distribution of Mahabharata

As shown in Fig. 4, most of the sentences have a length between 20 and 70 words. And very few sentences are having a length of less than 20, and some outliers do have higher lengths like the sentence on the 121,306 index is having a length of 1850 words.The text is not only divided into chunks of sentences but also into unit words to add more granularity into text preprocessing. This is achieved using the technique called word tokenization. Here, we used the 'natural language toolkit - word-tokenize()' method which divides the whole Mahabharata text into 2,749,461 uni-grams (only one word).

3.4 Text Normalization

The human written text includes function words and content words. These text data specifically the fictional text is a combination of all the grammatical ups and downs. Thus, these data do have high randomness. To reduce the randomness of the text and maintain the significant meaning of the text, the text normalization can be performed on the whole text.

On the Mahabharata text, we are applying two popular techniques stemming and lemmatization. These tasks are followed in the NLP pipeline to transform the fictional text into the standard form of the language. Both these tasks are followed by 'removing stop words' on the text.

In Mahabharata text, many words do not have critical significance but are used with high frequency throughout the whole epic to form the correct grammar. These words are not useful to improve the performance of any language model, and they will also take some computational time in the further analysis process. These words do not have any information in terms of sentiment analysis as well. So, it is advisable to remove stop words (words like a, an, the, are, have, etc.) along with the text normalization tasks.

Stemming with Stop Words: One word having the same semantic meaning can be written in many formats with human language. Stemming is a technique that removes the affixes and suffixes attached to the word and tries to bring out the stem word or root word from the text. Among popular stemming techniques like Lancaster stemmer, Porter Stemmer, and Snowball stemmer, we used Porter stemmer to get the root words of the whole Mahabharata text.

Lemmatisation with stop words: The process of Lemmatisation is designed with the same purpose which is addressed by stemming. Lemmatisation is also used to cutting down the words to their root word. However, in lemmatisation, the inflection of the word is not just broken-off, but it uses the concept of lexical knowledge. Using this converts, the words into base form. Thus, it holds the sentiments of the text more strongly. Here, we used 'wordnetlemmatizer' to achieve this task (Fig. 5).

```
# Original Sentence
sentences[470]

'the endeavours of duryodhana to engage yudhishthira again in the game; a
nd the exile of the defeated yudhishthira with his brothers.'

# Sentence after Stemming
sentences_stm[470]

'endeavour duryodhana engag yudhishthira game ; exil defeat yudhishthira
brother .'

# Sentence after lemmatisation
sentences_lemma[470]

'endeavour duryodhana engage yudhishthira game ; exile defeated yudhishth
ira brother .'
```

Fig. 5 Result of stemming and lemmatisation on Mahabharata Corpus

The selection between stemming and lemmatisation can be done based on the database on which the language model is going to be built. The Mahabharata is a fictional text, and to extract features from this large epic, a strong sentimental hold on the text is required. Thus, based on the comparison of stemming and lemmatisation, we decide to build the language model on lemmatized text.

3.5 The Language Model

The second objective of this paper is to find similar words from the Mahabharata fictional text. Here, we have a large amount of fictional text which can be considered as unannotated data for training a model. Thus, according to [17], Word2Vec is well liked model to be applied on data which do not have any adulteration. Word2Vec is a combination of two different algorithms applied together on corpus. These two algorithms are known as continuous bag of words (CBOW) and skip-gram. This model is developed with three different layers such as (a) input layer, (b) single hidden layer, and (c) output layer. The input layer is consisted with set of neurons which is having shape of the total number of words in the vocabulary. This vocabulary is specifically built according to the corpus. In this paper, our corpus is the book 'Mahabharata', and the vocabulary created (Fig. 6) based on this text is containing 25794 words.

The magnitude of a single hidden layer is equal to the dimensionality of the result word vector. Here, we trained a Word2Vec model to get the 100-dimension resultant vector. So, the size of the hidden layer is 100 dimensional. And the output layer is having the same magnitude as the input layer. Considering 'V' words in the vocabulary (where $V = 25{,}794$) and 'N' is the dimension of the resultant vector (where $N = 100$). Thus, the connections from the input layer to the hidden layer can be constituted by the WI matrix having the shape of V $\times$ N. Here, each row and column represents each word of vocabulary and the dimension of the resultant vector, respectively. Likewise, the connections from the hidden layer to the output layer can be constituted by a WO matrix having the shape of $N \times V$. Here, each

Fig. 6 Vocabulary based on Mahabharata corpus

Index	Vocab Words	Index	Vocab Words
4782	behaves	29391	palala
32282	arthakaman	14122	diffusing
25137	anangangahara	25456	bhuyas
14308	bread	15506	pravarakarna
2529	fulfilled	13356	ananda
9092	overspreads	24286	urdhvavahu
10737	vacant	19381	buyest
22088	durvaranah	5034	secretion
4901	befallen	8882	brahma-weapon
28502	dipaka	28471	well-instructed

```
sentences_lemma[3671]

'one day wait upon wrathful ascetic rigid vow durvasa name acquainted tru
th fully conversant mystery religion .'

sentences_lemma[3672]

'pritha possible care gratified wrathful rishi soul complete control .'

sentences_lemma[3673]

'holy one gratified attention bestowed maiden told satisfied fortunate on
e thee !'
```

Fig. 7 Sample corpus from the Mahabharata

row and column represent the dimension of the resultant vector and each word of vocabulary, respectively.

Considering the above sample corpus (Fig. 7), the vocabulary created based on this corpus can be represented as follows:

$Vocabulary_S = `one' : 0, `day' : 1, `wait' : 2, `upon' : 3, `wrathful' : 4, `ascetic' : 5, `rigid' : 6, `vow' : 7, `durvasa' : 8, `name' : 9, `acquainted' : 10, `truth' : 11, `fully' : 12, `conversant' : 13, `mystery' : 14, `religion' : 15, `pritha' : 16, `possible' : 17, `care' : 18, `gratified' : 19, `rishi' : 20, `soul' : 21, `complete' : 22, `control' : 23, `holy' : 24, `attention' : 25, `bestowed' : 26, `maiden' : 27, `told' : 28, `satisfied' : 29, `fortunate' : 30, `thee' : 31, `!' : 32$

The sample corpus vocabulary has 33 words. This vocabulary is considering each unique word given in the sample corpus. So, there are 33 input neurons and 33 output neurons. We have 100 neurons in the hidden layer. Thus, our connections neurons between the input layer to the hidden layer can be represented as WI(33 $\times$100), and the connection neurons between the hidden layer to the output layer can be represented as WO(100 $\times$ 33). Now before we train the Word2Vec model, these matrices are initialized with small random numbers.

Now looking at the corpus, if we want that Word2Vec model finds the relationship between the words ‘durvasa’ and ‘vow’; the word ‘durvasa’ is known as context, and ‘vow’ is known as the target.

Now, these inputs can be multiplied with the randomly initialized WI(33 $\times$100) matrix tending toward the hidden layer, and then, the output at the hidden layer will be multiplied with WO(100 $\times$ 33)matrix while tending toward the output layer. The target of this model is to compute probabilities for words at the output layer. This is achieved in Word2Vec as it implements the softmax function.

The idea behind using Word2Vec is, this model is used to represent the words by a vector of numbers. In our case, we provide the target word as input to the model. It will compute the cosign similarities between all other words available in the vocabulary and send it back as output with *top n* words.

4 Results

After text preprocessing, we applied Word2Vec on the corpus with 25,794 length vocabulary. We considered 'Hastinapur(location)', 'Arjuna (protagonist)', 'Gandiva (object)', 'Shakuni (antagonist)', 'dice (object)', 'Krishna (protagonist)', 'Siva (divine character)' and Panchali (female protagonist) as target words. These words are selected based on the popularity of the protagonist, location, and object covered in Mahabharata. The sample target words with their similar words along with the cosign similarity between target word and context word are shown in Table 1. *(Here, we are considering top five similar words)*

Table 1 Cosine similarity between words in Mahabharata

Context word	Cosine similarity	Context word	Cosine similarity
Sample target word: Hastinapur (Location)		Sample target word: Arjuna (Protagonist)	
Sheltered	0.8325892090797424	Partha	0.8605594038963318
Vaivaswata	0.8262888193130493	Dhananjaya	0.8489114046096802
Ansu	0.8186947107315063	Karna	0.8209618926048279
Coronet	0.8173437714576721	Vibhatsu	0.7797690629959106
Kurujangala	0.8155612349510193	Phalguna	0.7688585519790649
Sample target word: Gandiva (Object)		Sample target word: Shakuni (antagonist)	
Discus	0.8680269122123718	Shikhandi	0.9232674241065979
Wielder	0.8528252840042114	Somadatta	0.9199751615524292
Mace	0.7967674732208252	Duhsasana	0.9096741676330566
Thunder-bolt	0.7912233471870422	Vikarna	0.9092869758605957
Trident	0.7857180833816528	Satanika	0.9051489233970642
Sample target word: dice (Object)		Sample target word: Krishna (protagonist)	
War	0.7768427133560181	Kesava	0.8743714690208435
Suyodhana	0.7763647437095642	Janardana	0.7793148756027222
Match	0.7677909135818481	Vasudeva	0.7790760397911072
Wretched	0.7589513659477234	Keshava	0.768439769744873
Gambling	0.7584301829338074	Vibhatsu	0.7160406708717346
Sample target word: Siva (protagonist - God)		Sample target word: Panchali (female protagonist)	
Sthanu	0.9016269445419312	Vabhru	0.9394650459289551
Three-eyed	0.900800883769989	Vibhishana	0.9367938041687012
Boon-giving	0.9007714986801147	Ahuka	0.932561993598938
Isana	0.8950022459030151	Sudeva	0.9305785298347473
Skanda	0.8801624774932861	Vrishaparvan	0.9267337322235107

5 Conclusion

In this paper, the NLP-based experiment on the Mahabharata is carried out on a basic level. We trained the Word2Vec model on the corpus to get the most similar words from the text itself. The reason behind selecting Word2Vec is to deal easily with the high-dimensional word vectors. In this paper, we can reach the basic aspects like uni-gram vocabulary, sentence distribution, 100-dimensional vector representation, and word similarities.

6 Future Scope

Though in the current scenario, we do have similar models like 'GloVe' and 'FastText' which we are targeting to apply and compare the results. The comparison of these models will bring a robust argument that which model is giving the best result on fictional text. Apart from the text-similarity, we are also targeting to perform noun extraction, word cloud for the various protagonist of Mahabharata, creating custom corpus, parts-of-speech-tagged word corpus, chunked phrase corpus, and create categorized text corpus of Mahabharata to make it open source for various experiments. Apart from these, we are targeting text classification on the Mahabharata. The major purpose to opt for these NLP techniques on Mahabharata is to understand and observe the behavior of the Mahabharata protagonist. These observations can be matched with human behavior with the help of a specific questionnaire based on organizational behavior. This can provide a profile of a person and production capacity in his/her working environment. Thus, the results of this paper can be mapped with future research to identify the professional perspective of a human personality based on the Mahabharata.

References

1. Ganguli KM (1886) Mahabharata of Krishna Dwaipayana Vyasa (English translation in the year of 1886). In: Roy PC (ed) Calcutta Bookseller
2. Koch D. Astronomical dating of the Mahābhārata war, Version 2. Gilgamesh
3. Sudiatmika IBK, Artana M, Utami NW, Putra MAP, Dewi EGA (2021) Mask R-CNN for Indonesian shadow puppet recognition and classification. J Phys Conf Ser 1783, 1 Feb 2021 (IOP Publishing). http://creativecommons.org/licenses/by/3.0/
4. Ghorai M, Purkait P, Santra S, Samanta S, Chanda B (2016) Bishnupur heritage image dataset (BHID). In: Proceedings of the tenth Indian conference on Computer Vision, Graphics and Image Processing—ICVGIP '16. In: The tenth Indian conference. ACM Press, New York
5. Rajarajan K, Kalidos R (2021) From purā?ic to folk: The 'Kirātārjunīyamballade' and visuals. In: Imago E (ed) Universidad Complutense de Madrid (UCM), vol 10, pp 355–366
6. Mohanty P, Mathura (2021) Orality, literacy and translation. Rupkatha J Interdiscip Stud Humanit 13, 2 (Aesthetics Media Services)

7. Chu CX, Razniewski S, Gerhard W (2019) TiFi: Taxonomy induction for fictional domains. In: The World Wide Web Conference on—WWW '19. ACM Press, New York, The World Wide Web conference, San Francisco, CA, USA
8. Chu CX, Razniewski S, Gerhard W (2020) ENTYFI. In: Proceedings of the 13th international conference on web search and data mining. ACM, New York, WSDM '20: The thirteenth ACM international conference on web search and data mining, Houston, TX, USA
9. Subhashree, Kumar Sreenivasa P (2018) Enriching domain ontologies using question-answer datasets. In: Proceedings of the ACM India joint international conference on data science and management of data. ACM, New York, CoDS-COMAD '18: The ACM India joint international conference on data science and management of data, Goa, India
10. Marcinkowska-Rosół M, Sellmer S (2021) The mind as container. In: Mnemosyne, Brill, pp 1–25
11. Post M, Callison-Burch C, Osborne M (2012) Constructing parallel corpora for six Indian languages via crowd sourcing. In: Proceedings of the 7th workshop on statistical machine translation, Montréal, Canada, 7–8 June 2012, pp 401–409. c2012 Association for Computational Linguistic
12. Das D (2016) A computational analysis of Mahabharata. In: International conference on natural language processing ICON-2016At. IIT-BHU, Varanasi, India
13. Ramachandran P, Sharma R (2013) Are you a Kaurava or a Pandava at work?: management lessons from the Mahabharata. IUP J Soft Ski VII, 2 (Oxford Univ Press)
14. Sivakumar KS (2018) Individual development in the light of the Bhagavad Gita
15. Iyengar RV (2013) Arjuna and the art of concentration. IUP J Soft Skill VII, 4
16. Debroy D (2018) The changing treatment of disability in children's literature in India. In: Bookbird, Project Muse, vol 56, issue 4, pp 14–21
17. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. In: arXiv preprint arXiv:1301.3781

A Real-Time Driver Drowsiness Detection Using OpenCV, DLib

Srinidhi Bajaj, Leena Panchal, Saloni Patil, Krutika Sanas, Harshita Bhatt, and Swapnali Dhakane

Abstract Every year thousands of people die around the world in motorway accidents, and one of the main reasons for this is drivers' drowsiness and fatigue. According to a survey by the Central Road Research Institute (CRRI) in 2019, drivers who exhaust themselves doze off, while driving are accountable for about 40% of road mishaps. To reduce the road mishaps, a system to monitor driver's alertness by detecting the visual features of the driver by finding the drowsiness state of the driver is proposed. It deals with an algorithm which considers the frequency of the eye-blink called PERCLOS, that make use of the eye coordinates obtained from Dlib's Haar cascade model to determine eye's state of the driver either open or close and sounds an alarm if the driver is found to be in drowsy state, the warning can be deactivated manually rather than automatically. This algorithm performs better than current drowsiness detection systems in both accuracy as well as speed at adequate lighting conditions. The frames captured of driver are of 640*480 resolution at over 20 fps to determine drowsiness of the driver and give accuracy of 98%. It is also affordable as it does not require any expensive hardware, only a built-in Android camera is required to provide a warning sound when the proposed system predicts that the driver is drowsy. This research result can serve as an important component

S. Bajaj (✉) · L. Panchal · S. Patil · K. Sanas · H. Bhatt · S. Dhakane
Department of Information Technology, MKSSS'S Cummins College of Engineering For Women, Pune, Maharashtra, India
e-mail: srinidhi.bajaj@cumminscollege.in

L. Panchal
e-mail: leena.sharma@cumminscollege.in

S. Patil
e-mail: saloni.patil@cumminscollege.in

K. Sanas
e-mail: krutika.sanas@cumminscollege.in

H. Bhatt
e-mail: harshita.bhatt@cumminscollege.in

S. Dhakane
e-mail: swapnali.dhakane@cumminscollege.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_64

in ADAS, and it can ensure safety of drivers and minimise financial and personal losses caused by accidents.

Keywords OpenCV · Haar cascade · Eye detection · Eye aspect ratio · Drowsiness detection · Android application

1 Introduction

A road accident can be the cause of untimely deaths for more than one reason. But one of the causes of traffic accidents is road mishaps. Many people die in traffic accidents each year from driving fatigue caused by insomnia, drunkenness, drug abuse, heat, or alcoholic beverages [1, 2]. A co-passenger will usually alert the driver to the driver's lack of attention. It is also possible for co-passengers to sleep while the driver drives alone. Every year, over 100,000 road accidents are caused by driving under the influence of drowsiness [3]. That is the equivalent of 2.2% of all road deaths [3]. It is estimated that 91,000 accidents involving drowsy drivers were reported to the National Highway Traffic Safety Administration in 2017. The reality may be more dangerous than the numbers suggest. It has become increasingly necessary over the course of the year as the number of vehicles grows worldwide. Sleep deprivation leads to drowsiness. During this period, the driver is unresponsive to the overall driving situation, resulting in a delay in reaction time to any unforeseen circumstances [4]. As for the behaviour, the driver may seem to have droopy eyes and eyelid closure. This could be dangerous in situations that demand constant focus. As the driver in a drowsy state is unaware of their conditions and dozing time, they have to be monitored from an outside environment. To tackle this problem caused by the drowsiness of the driver, a lot of research is done to solve this issue. Existing methods to detect drowsiness are using machine learning techniques like classification, regression, deep learning, neural networks, etc. The existing implemented algorithms are either from one of the neural network categories or in the computer vision-based classifier category. But, most of these technologies are patented to high-end vehicles only. These sleep detection procedures can be classified depending on a number of factors such as motor, behavioural, and physical factors.

1.1 Physiological Level Approach

This approach is an intrusive method, wherein sensors are used to get the pulse rate, heart rate, and brain activity of the driver. ECG is used to calculate the variations in heart rate and also detect other conditions for drowsiness. The different signals such as ECG (electrocardiogram), EEG (electroencephalogram), and EMG (electromyogram) are evaluated, and then it is decided whether the person driving the vehicle is drowsy or not [5].

1.2 Behavioural-Based Approach

In this particular approach, the frequency of eye blinking, position of head, yawning, etc., of the driver is monitored and captured through a camera placed in the front, and the driver is alerted if any of these drowsiness symptoms are detected by the algorithms applied [6].

The objective of the proposed system is to implement surveillance for the driver with the help of the camera in smartphone to track down the signs of drowsiness, alerting the driver regarding the same, and storing the performance of the driver. The algorithm analyses visual features of the driver, specifically eyes to detect fatigue in the driver while driving [7]. The model proposed is a non-intrusive system which cannot distract the driver while driving. It uses an image processing method based on OpenCV and Dlib in the Android app which captures the photos of the driver at a rate of 20 fps. If the application finds that the driver is drowsy after the eyelid closure is 2–5 times, it activates the alarm to alert the driver.

2 Related Work

Several approaches have been proposed to measure the state of drowsiness while driving. According to driver drowsiness detection using MATLAB, photograph acquisition, photograph processing and computer imaginative, and prescient toolbox to discover whether eyes are closed or not using the included webcam of pc, and threshold of 70% is decided for distinguishing closed eye and open eye [8]. In this paper with the help of DriCare, yawing, duration of blinking, and time for which eyes are closed were detected for the driver's drowsiness status. In the proposed system, both the eyes and mouth features were captured and if the driver is found to be drowsy, the system could alert the driver by a fatigue alarm. An accuracy of 92% is achieved by the proposed model [9]. Another approach detects the drowsiness of the driver with the help of the variation in heart rate. Accuracy of output for the logistic regression test is over 92% [10].

Facial regions of detection are defined by facial key points. As the CNN architecture, VGG16 is used. Android app captures the real-time identified applicable sponsor/s here. VGG16 has acquired the highest accuracy, and the training accuracy acquired is 91% [11]. This method is implemented using Internet of Things technology. It has used machine learning, openCV-eyelocator, IR sensors, and the GPS service. The advantages of the system are the system with cameras focusing on user hash-free user assistance provided, it is communication systems, also alcohol discovery framework in vehicles gives a framework to vehicles and to different vehicles also [12]. Another approach for real-time drowsiness detection is implemented using a deep neural network to an Android application. The proposed model achieves accuracy of 80% but still needs space for better improvement [13].

3 Technologies Used

3.1 OpenCV

OpenCV is an open source computer vision library designed for image processing that can be used to perform tasks like face detection, object tracking, landmark detection, etc. The OpenCV and Dlib support the Android platform to detect the faces from the live frame by predicting facial features for eye location and detects the driver's drowsiness [14]. OpenCV supports Haar cascade-based object detection. Haar cascades are machine learning-based classifiers that calculate different features like edges, lines, etc., in the image.

3.2 Dlib

Dlib is a C++ toolkit that contains machine learning algorithms and tools used in the development of software to resolve real-world issues [15]. It is helpful for detecting objects in images which includes frontal face detection and object pose estimation. It provides high-quality face recognition.

3.3 Android Application

The Android application is the main HCI of the model. Application is developed using Android Studio, and Module SDK is Android API 30 platform [16]. App will record the real-time video of the driver, and based on the input video, it will sound an alarm to alert the driver or wait for the next frame where the driver is found to be drowsy. It is secure as the app cannot access parts of the system for which it is not given permission. The application only accesses the camera and location of the driver.

4 Methodology

There are seven steps in the proposed algorithm as described in Fig. 1. Tests were conducted using the driver's mobile camera as the video feed.

Figure 1 shows the operation of a proposal system that uses the EAR to calculate whether eyes are open or closed.

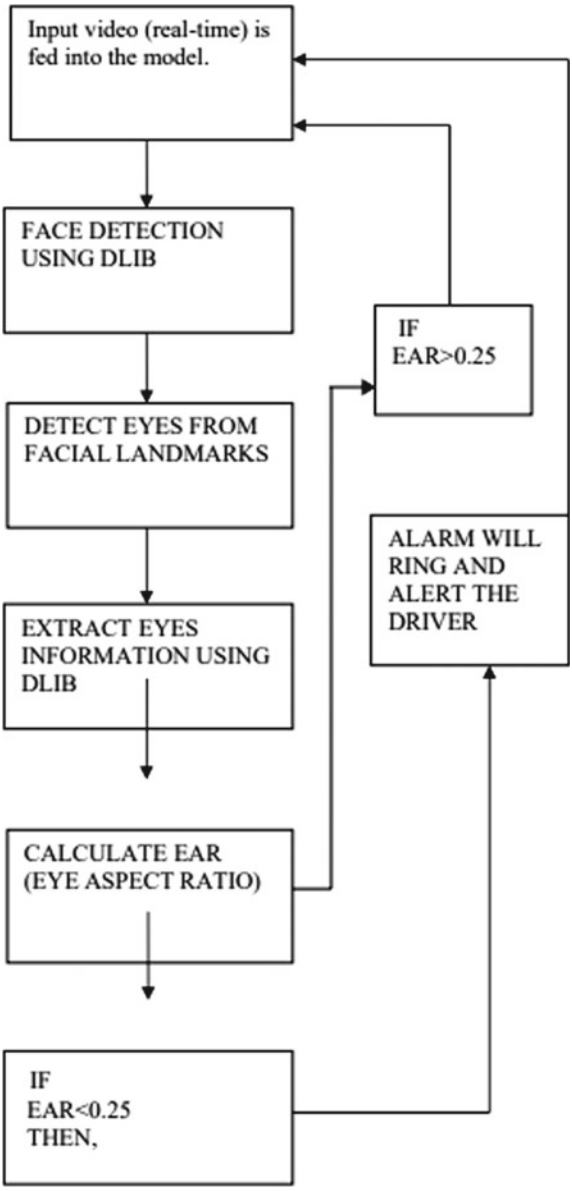


Fig. 1 Flow chart of the drowsiness detection

4.1 Pre-processing

Camera is placed in front of the driver's face and the video is being captured. The processing speed of the received video is 20 fps. It then examines this frame and converts it to grayscale.

4.2 Face Detection

In addition to heights, weights, and facial features, Haar-based classifiers include other features as well. The classifier uses both positive and negative samples for the classification. Features based on positive dimensions are extracted [17]. Edge detection based on Haar detection is applied. The output of the edge detector is stored in an array. Detecting the driver's face is achieved by extracting the eye features and applying parallel processing. The pixels of the upper area of the face are only considered because just a few pixels below the upper edge of the face are the eyes. We apply edge detection to the area we marked previously.

4.3 Facial Mapping Using Dlib

This algorithm uses the Dlib for detecting and isolating facial features by using the pre-trained Dlib facial features detectors [18]. As shown in Fig. 2, a human face is evaluated and displayed as a face point using 68 Cartesian coordinates.

Fig. 2 Facial features prediction through Dlib and Opencv. *Source* MEDIUM, using face emotion recognition model with dlib and openCV in iOS application

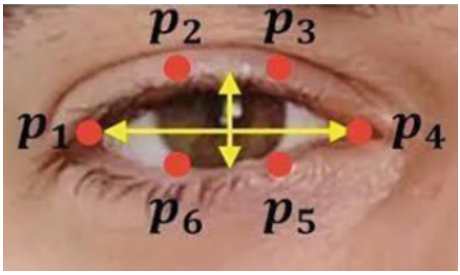


Fig. 3 EAR and drowsiness detection using Dlib. *Source* MEDIUM, how to detect eye blinking in videos using dlib and OpenCV in Python

A 68-point iBUG 300 database has been used for training these Dlib facial landmark predictions.

4.4 Eye Closure Detection

A measure of the eye aspect ratio (EAR) indicates whether the eye is open or closed by determining its state. Face landmarks plotted by Python's Dlib library can be used for calculating this. EAR is computed by taking facial landmark points from the Dlib library, as shown in the following figure.

In Fig. 3, P_1, P_2, P_3, P_4, P_5, and P_6 are the pupil coordinates. When eyes are open, EAR is near 0.25. If the EAR is less than 0.25, a conclusion is drawn about drowsiness in a person.

The eye aspect ratio (EAR) is calculated for both eyes using,

$$\frac{(|P_2 - P_6| + |P_3 - P_5|)}{2(|P_1 - P_4|)} \tag{1}$$

The numerator is the vertical distance between the upper and lower eyelids. The denominator represents the horizontal distance between both the upper and lower eyelids. Calculate the average of the EAR values obtained from each eye. The distance between both eyelids decreases when a person is drowsy. The frequency of the eye movement is fixed [19]. The screen displays an alarm when the value exceeds the threshold. An alert will appear if the eye aspect ratio is less than 0.25 in 20 consecutive frames.

Additionally, it calculates the ratio of drowsy frames to total frames over time.

PERCLOS can be calculated by using the following formula given below:

$$\text{PERCLOS} = \text{NDrowsy}/\text{NTotal} \times 100\% \tag{2}$$

where NDrowsy is the number of frames in which a driver is found to be drowsy as judged by classifier and NTotal is the total number of frames in a specific time. If PERCLOS > Threshold, the driver is said to be drowsy.

5 Experimental Results

The system proposed can detect the drowsiness state of the driver irrespective of any obstruction of glasses, hearing aids, etc.

As shown in Fig. 4, the proposed system detects if the driver's eye is open or closed based on calculated EAR.

As shown in Fig. 5, if the driver is found to be drowsy, an alert box will appear along with a warning sound. The experimental results of the proposed system are summarised as (Table 1).

The above-mentioned conditions state that after the detection of eyes if found closed, then the alarm would beep, else it would continue to detect (Table 2).

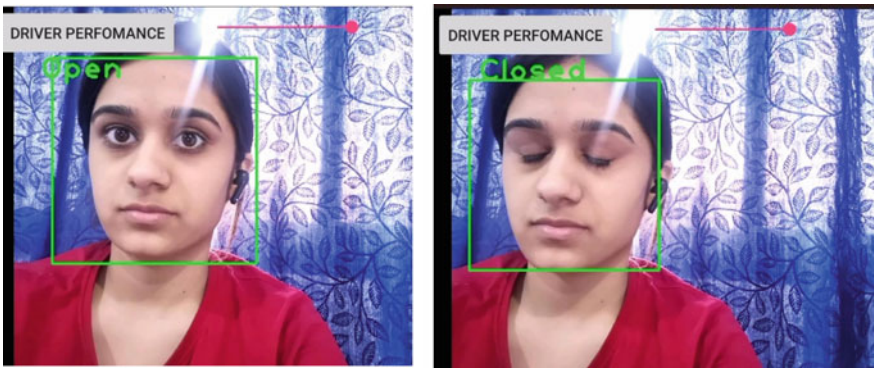


Fig. 4 Eye posture detection of the driver

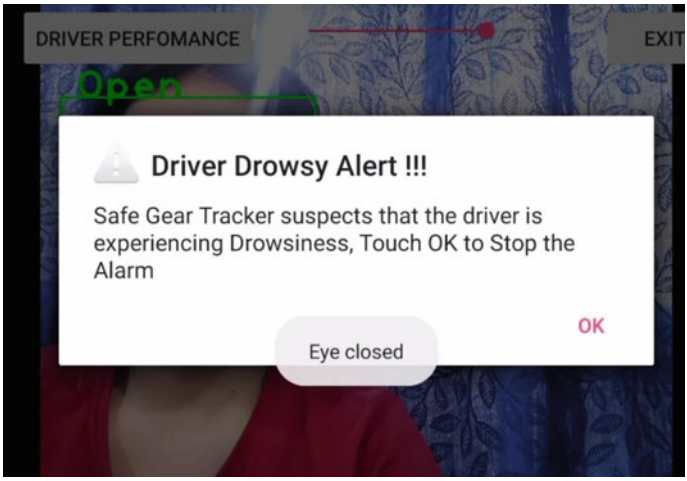


Fig. 5 Proposed system alert the driver

Table 1 Test cases that were conducted for drowsiness detection of the driver

Test cases	Eyes detected	Eye closure	Result
1	No	No	None
2	Yes	No	None
3	Yes	Yes	Voice alert

Table 2 Test cases that were conducted on different individuals in different conditions

Serial No	Conditions	Accuracy of eye detection (%)	Accuracy of drowsiness detection (%)
Individual 1	Day time	100	100
Individual 2	Placed at a distance from right side	99	98
Individual 3	Night time	96	98
Individual 4	Placed at a distance from left side	94	96
Total		97.25	99.25

$$\text{Accuracy of detection of eye} = (\text{Total number of times eyes detected in frame}) / (\text{Total number of times eyes detected in frame} + \text{Total number of times eyes not detected in frame})$$

$$\text{Accuracy of detection of drowsiness state} = (\text{Total number of times alarm activated}) / (\text{Total number of times alarm activated} + \text{Total number of times alarm not activated}).$$

According to the implementation of the system, if the eyes remain closed for more than the value set as threshold for the number of frames, then the driver is drowsy.

6 Conclusion

A non-intrusive system to locate the eyes and monitor driver drowsiness was developed. Information about the eye's position is obtained through OpenCv and Dlib algorithms. The system is able to detect whether the eyes of the driver are closed or open. If the eyes are detected to be closed, a warning alarm is triggered. Those drowsiness statuses are lively updated to the owner in the website. Image processing proves to be accurate and reliable in detection of drowsiness. In addition to the proposed work, absence of proper lighting, detection of the facial features of the

driver like eyes will be hard. So, the system will take time to detect eyes in dim light. The system can detect only one face at a time. If there is more than one face in the frame, the results may differ. If the face is tilted during face detection, the 68-point detection will perceive the face in the same manner, and it is not a major problem as after this only eyes are considered as ROI.

References

1. Dawson D, Reid K (1997) Fatigue, alcohol and performance impairment. Nature 388:235–235. https://doi.org/10.1038/40775
2. Williamson AM, Feyer AM, Mattick RP et al (2001) Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. Accid Anal Prev 33:313–326. https://doi.org/10.1016/s0001-4575(00)00045-2
3. Zhang G, Yau KKW, Zhang X, Li Y (2016) Traffic accidents involving fatigue driving and their extent of casualties. Accid Anal Prev 87:34–42. https://doi.org/10.1016/j.aap.2015.10.033
4. Dinges DF (1995) An overview of sleepiness and accidents. J Sleep Res 4:4–14.https://doi.org/10.1111/j.1365-2869.1995.tb00220.x
5. Sommer D, Golz M (2010) Evaluation of PERCLOS based current fatigue monitoring technologies. In: 2010 annual international conference of the IEEE engineering in medicine and biology. IEEE, pp 4456–4459
6. Gao Y, Wang C (2017) Fatigue state detection from multi-feature of eyes. In: 2017 4th international conference on systems and informatics (ICSAI). IEEE, pp 177–181
7. Ma X, Chau L-P, Yap K-H (2017) Depth video-based two-stream convolutional neural networks for driver fatigue detection. In: 2017 international conference on orange technologies (ICOT). IEEE, pp 155–158
8. Jabbar R, Al-Khalifa K, Kharbeche M et al (2018) Real-time driver drowsiness detection for android application using deep neural networks techniques. Procedia Comput Sci 130:400–407. https://doi.org/10.1016/j.procs.2018.04.060
9. Biswal AK, Singh D, Pattanayak BK et al (2021) IoT-based smart alert system for drowsy driver detection. Wirel Commun Mob Comput 2021:1–13. https://doi.org/10.1155/2021/6627217
10. Sinha A, Aneesh RP, Gopal SK (2021) Drowsiness detection system using deep learning. In: 2021 seventh international conference on bio signals, images, and instrumentation (ICBSII). IEEE, pp 1–6
11. Bajaj P, Ray R, Shedge S et al (2021) Synchronous system for driver drowsiness detection using convolutional neural network, computer vision and android technology. In: 2021 7th international conference on advanced computing and communication systems (ICACCS). IEEE, pp 340–346
12. Jain M, Bhagerathi B, C N DrS (2021) Real-time driver drowsiness detection using computer vision. Int J Eng Adv Technol 11:109–113. https://doi.org/10.35940/ijeat.A3159.1011121
13. Sabu S (2021) Driver drowsiness detection and warning system. Int J Res Appl Sci Eng Technol 9:657–660. https://doi.org/10.22214/ijraset.2021.34246
14. Vishnu R, Vishvaragul S, Srihari P et al (2021) Driver drowsiness detection system with opencv and keras. J Phys Conf Ser 1916:012172. https://doi.org/10.1088/1742-6596/1916/1/012172
15. Mohanty S, Hegde SV, Prasad S, Manikandan J (2019) Design of real-time drowsiness detection system using Dlib. In: 2019 IEEE international WIE conference on electrical and computer engineering (WIECON-ECE). IEEE, pp 1–4
16. Galarza EE, Egas FD, Silva FM et al (2018) Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone, pp 563–572

17. Magán E, Sesmero MP, Alonso-Weber JM, Sanchis A (2022) Driver drowsiness detection by applying deep learning techniques to sequences of images. Appl Sci (Switzerland) 12.https://doi.org/10.3390/app12031145
18. Rahman A, Sirshar M, Khan A (2015) Real time drowsiness detection using eye blink monitoring. In: 2015 national software engineering conference (NSEC). IEEE, pp 1–7
19. Iyer V, Vanjari A, Patil V et al (2021) Driver drowsiness detection system. 10

Operational Availability Optimization of Cooling Tower of Thermal Power Plants Using Swarm Intelligence-Based Metaheuristic Algorithms

Ashish Kumar, Deepak Sinwar, Vijaypal Singh Dhaka, and Sunil Kr. Maakar

Abstract Cooling towers are mainly utilized to disperse the heat of thermal power plants (TPP). The availability of cooling tower is directly proportional to the maximum availability of TPP. To ensure the maximum availability of cooling towers, a mathematical model is developed followed by optimization using four swarm intelligence-based metaheuristic algorithms, viz. Grey Wolf Optimizer, Grasshopper Optimization Algorithm, Dragonfly Algorithm, and Whale Optimization Algorithm. The Markovian birth–death process and Chapman-Kolmogorov differential–difference equations are utilized to derive the objective function of availability associated with the proposed model. It is observed from the numerical investigation that the Whale Optimization Algorithm performs better than all other metaheuristic algorithms in providing the optimized values of various failure and repair rates and predicting the overall availability of the cooling tower.

Keywords Grey Wolf Optimization · Grasshopper Algorithm · Dragonfly Algorithm · Whale Optimization Algorithm · Reliability · Availability · Optimization

1 Introduction

Optimizing the availability of process industries or subsystems is one of the challenging areas for enhanced reliability. Many researchers carried out several studies on the modelling of the industrial systems, viz. thermal power plants, fertilizers plant, paper industries, shoe manufacturing, etc. to assess their performances. Cooling

A. Kumar
Department of Mathematics and Statistics, Manipal University Jaipur, Jaipur, India

D. Sinwar (✉) · V. S. Dhaka
Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India
e-mail: deepak.sinwar@gmail.com

S. Kr. Maakar
School of Computing Science and Engineering, Galgotias University, Greater Noida, India

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_65

tower of thermal power plants is one such system comprises six units. The primary role of cooling tower is to disperse the heat from the routine operations of a thermal plant. Plenty of work has been carried out to assess and enhance the availability of similar systems. Khosravi et al. [1] presented a new method for evaluating the reliability of power plants. Souza et al. [2] performed reliability investigation for a thermal turbine and gas power plant. On the other hand, a stochastic model to assess the impact of environmental conditions on the reliability measures is developed by Saini and Kumar [3]. Gupta et al. [4] applied Markovian methodology to investigate the availability of generators of steam turbine power plants. They utilized reliability investigation measures, namely reliability, availability, maintainability, and durability (RAMD) are utilized for this purpose. Similarly, RAMD analysis to evaluate the maintainability and reliability of generator is presented by Gupta et al. [5]. All these approaches are only concentrated on obtaining the local solution of system measures. But since the last few decades, optimization of the performance measures of industrial systems attracted the concentration of the researchers. For predicting the performance of cooling towers, Khodakaram-Tafti and Golneshan [6] developed a mathematical model followed by extension to three aligned towers. On the other hand, for parameter optimization of solar cells, Sharma et al. [7] developed an algorithm based on opposition tunicate swarm approach. In addition to the development of mathematical models for reliability assessments, the usage of metaheuristic approaches is also gaining good popularity. In this regard, Jagtap et al. [8] investigated the performance of turbogenerator followed by optimization of availability using particle swarm optimization (PSO). Zhang [9] conducted the transient performance evaluation of reliability measures using PSO. Similarly, the availability optimizations of condenser units is carried out by Saini et al. [10]. They first presented a stochastic model followed by parameter optimization process using genetic algorithm (GA) and PSO. Recently, Kumar et al. [11] utilized two well-known metaheuristic approaches, i.e. GA and PSO for optimizing the availability of cooling towers. They first presented a stochastic model of the system and then computed the optimized values of various parameters of different states. Out of several types, swarm intelligence (SI)-based metaheuristics are gaining popularity due to their numerous advantages. A very few literatures are observed on optimizing the availability computations using such approaches. Considering all such evidences, a comparative study of four SI-based approaches, viz. Grey Wolf Optimizer (GWO) [12], Grasshopper Optimization Algorithm (GOA) [13], Dragonfly Algorithm (DA) [14], and Whale Optimization Algorithm (WOA) [15] is carried out to optimize the availability of the cooling towers of thermal power plants by using swarm intelligence-based algorithms. For this purpose, mathematical model of cooling tower proposed by Kumar et al. [11] is utilized. The Markovian birth–death process and Chapman-Kolmogorov differential–difference equations are utilized to derive the objective function of availability associated with the proposed model. It is observed from the numerical investigation that WOA outperforms all other algorithms in predicting the availability of the cooling tower.

The rest of the paper is organized as follows: Sect. 2 presents mathematical model development and optimization strategies applied in this work; numerical analysis is presented in Sect. 3. Finally, Sect. 4 presents concluding remarks.

2 Materials and Methods

2.1 Model Development

In this section, the mathematical model of the cooling tower is developed under the assumptions and notations given by Kumar et al. [11]. An exponential distribution is followed by failure and repair rates, whereas all associated random variables are statistically independent. The recurrence relations of the mathematical model are described as follows:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + u_0\right) N_o(t) = \sum_{i=1}^{7} \sigma_i N_i(t) \tag{1}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + \sigma_k\right) N_k(t) = \vartheta_k N_0(t) \quad \text{where } k = 1, 2, 3, 5, 6, 7 \tag{2}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + u_0 + \sigma_4\right) N_4(t) = \vartheta_4 N_0(t) + \sum_{s=1}^{7} \sigma_s N_{s+7}(t) \tag{3}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + \sigma_l\right) N_l(t) = \vartheta_l N_4(t) \quad \text{where } l = 1, 2, \ldots, 7 \tag{4}$$

Using the normalization criteria that $\sum_{i=0}^{14} N_i(t) = 1$, we get

$$\Rightarrow N_0 = [1 + \left(1 + \frac{\vartheta_4}{\sigma_4}\right)\left(\sum_{i=1}^{7} \frac{\vartheta_i}{\sigma_i}\right]^{-1} \tag{5}$$

Steady state availability is derived as follows using above set of equations:

$$\Rightarrow A_v = \left(1 + \frac{\vartheta_4}{\sigma_4}\right) N_0 \tag{6}$$

Equation (6) is treated as the objective function, and all the failure and repair rate parameters are the decision variables. Here, Eq. (6) is optimized using SI-based approaches (GWO, GOA, DA, and WOA).

2.2 Optimization Strategies

Optimization of a problem can be either minimization or maximization of an objective function. It deals with finding the best values of variables of that problem. As discussed earlier, for optimizing the failure rates, repair rates and overall availability of cooling tower of thermal power plant, four different population-based metaheuristic approaches are utilized. All these algorithms are population-based metaheuristic approaches that are also known as swarm intelligence approaches. Population-based metaheuristic approaches have several advantages as compared to a single solution-based approaches [12]: various solutions of a search space from a collaborative effort; each solution helps others for avoiding local optimal solutions; and better exploration capabilities due to collaborative intelligence. This section will brief about the working principle of all these algorithms.

2.2.1 Grey Wolf Optimizer (GWO)

Mirjalili et al. [12] established GWO in 2013 inspired by hunting mechanism of Grey Wolves. It works on the principle of leadership behaviour of wolves. In GWO, wolves are categorized into four different classes, viz. alpha, beta, delta, and omega. In addition, they employed three different ways of hunting: searching, encircling, and attacking a prey. Generally, grey wolves prefer living in a group (approx. 5–12 in numbers). In this group, alpha wolves act as decision-makers because they are the leaders of the group. On the other hand, beta wolves are subordinates/advisors of the alpha for assisting in the decision-making process. Also, in the absence of alpha wolves, the beta wolves can be found as the best candidate to serve the purpose. Omega is the lowest ranked wolf in the group because they must submit to all others in the group. In case a wolf is not amongst these, then it is treated as delta. Delta must submit to first two types of wolves, but they can overlook the third one (i.e. omega). The pseudocode of the GWO is mentioned as follows:

1. Population initialization of wolves.
2. Initialization of random parameters (a, A, and C).
3. Fitness computations of each search agent (SA) (X_α: best agent, X_β: second-best agent, X_δ: third-best agent).
4. Repeat till maximum iterations or satisfaction of stopping criteria:
   - For each SA
     - Update the position of SA based on search criterion
   - Update random parameters
   - Calculate fitness of all SA
   - Update the values of each SA
5. Return X_α.

2.2.2 Grasshopper Optimization Algorithm (GOA)

Saremi et al. [13] developed GOA in 2016 inspired by the behaviour of grasshoppers. Grasshoppers are one of the largest swarms of this creature. In their lifetime, nymph grasshoppers jump and move like rolling cylinders. During their migration from one place to another, they eat various types of vegetation, become adult, and develop swarm in the air. Grasshoppers have three main characteristics, viz. slow movement during larval phase, rapid and long-range movements during adulthood, and seeking for food. Mathematically, the grasshopper optimization algorithm can be presented as depicted in Eq. (7)

$$X_i = S_i + G_i + A_i. \tag{7}$$

Here, X_i indicates the position of ith grasshopper, S_i as the social interaction, G_i as gravity force, and A_i as wind advection. Equation (7) can be extended to Eq. (8) to include the randomness in the behaviour of grasshoppers as follows:

$$X_i = r_1 S_i + r_2 G_i + r_3 A_i. \tag{8}$$

Here, r_1, r_2, and r_3 indicate the random numbers in a range of 0 and 1. In addition, $d_{i,j}$ is a function indicating the distance between i and j. Function S represents the social forces based on d as indicated in Eq. (9):

$$S_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} s\left(d_{ij}\right)\widehat{d_{ij}} \tag{9}$$

2.2.3 Dragonfly Algorithm (DA)

Mirjalili [14] developed a new SI-based algorithm called Dragonfly Algorithm in 2016. The basis of DA lies in static (hunting) and dynamic (migration) behaviour of dragonflies. Static swarm is a kind of small group for hunting other flying preys, whereas dynamic swam (large number of dragonflies) is for migrating over long distances. The basic idea of the algorithm is mentioned in the form of pseudocode as follows:

1. Initialization of swarm (population of dragonflies).
2. Initialization of step vectors $\Delta X_i (i = 1, 2, \ldots, n)$.
3. ***Repeat*** till maximum iterations or satisfaction of stopping criteria:
    - Computation of objective values of each particle in the swarm
    - Update sources of food, enemies, radius, and other parameters

- ***If*** there is at least one neighboring dragonfly, then update velocity and position vectors ***else*** update position vector only
- Update new position using boundary values of variables

4. End

2.2.4 Whale Optimization Algorithm (WOA)

Whale are one of the extreme intelligent animals with plenty of emotions. Mirjalili and Lewis [15] developed a new SI-based metaheuristic algorithm called Whale Optimization Algorithm in 2015. Generally, whales are found in groups. However, they can live alone as well. To exploit, they encircle prey and form a spiral bubble-net. The basic mechanism of WOA is presented in the form of pseudocode as follows. For detailed description, please refer to [15].

1. Initialization of swarm of whales (search agent) and X to be the best search agent (SA).
2. Fitness computations of each SA.
3. ***Repeat*** till maximum iterations or satisfaction of stopping criteria.
   - ***For*** each SA update the values of parameters (a, A, C, l and p):
     - ***If*** $p < 0.5$ and

       If ($|A| < 1$) then update the position of current SA
       otherwise select a random SA and update the position of current SA
     - Otherwise (if $p \geq 0.5$) update the value of current SA
   - If any better solution is found, then update X
4. Return X and End

3 Numerical Results and Discussion

As mentioned in the previous section, to estimate the influence of failure and repair rates on the overall availability of cooling tower, the Eq. (6) is optimized using four SI-based metaheuristic algorithms. The initial parametric values of various failure (F) and repair (R) rates are mentioned as follows:

$$F_1 : 0.0002 - 0.08;\ F_2 : 0.0001 - 0.065;\ F_3 : 0.00002 - 0.0095;$$
$$F_4 : 0.00001 - 0.008;\ F_5 : 0.00025 - 0.085;\ F_6 : 0.0007 - 0.5;$$
$$F_7 : 0.000035 - 0.007;\ R_1 : 0.009 - 1.2;\ R_2 : 0.0075 - 1.15;$$
$$R_3 : 0.005 - 0.95;\ R_4 : 0.003 - 0.8;\ R_5 : 0.007 - 1.05;$$
$$R_6 : 0.0055 - 0.9;\ R_7 : 0.0048 - 1.025$$

For simulating the experimental evaluation, we have utilized RStudio (version 1.2.5042) on Windows10 64-bit with 8 GB of primary memory and Intel Core i7 7th generation CPU. A series of experiments were performed, and the overall availability of 99.87% is achieved with the help of WOA as given in Tables 1 and 3. Table 1 gives the values of availability obtained by different algorithms with a population size of 40 on 500 iterations, whereas Table 3 gives the same with a population of 100 on 1000 iterations. As given in Tables 1 and 3, an extreme impact of population and iteration was not observed in these experiments. Tables 2 and 4 represent the optimized values of failure and repair rates during the availability computations .

Figures 1 and 2 give the optimized values of availability function (Eq. 6) as computed by four SI-based metaheuristic approaches. It can also be observed from Tables 1 and 2 that the WOA outperformed other approaches in terms of maximum availability and time efficiency (very less elapsed time amongst all).

Table 1 Optimum value of availability using different metaheuristic algorithms on a swarm 40 with 500 iterations

Algorithm	Population	Iterations	Availability	Elapsed time (s)
GWO	40	500	0.9986769	4.08
GOA	40	500	0.9859904	33.37
DA	40	500	0.9140052	7.89
WOA	40	500	**0.9986771**	**0.93**

Table 2 Optimized values of failure and repair rates of the cooling tower using different SI-based algorithms (N_Pop: 40, N_Iter: 500)

Optimized value of failure and repair rates	GWO	GOA	DA	WOA
F_1	0.000200	0.000200	0.029351	0.000200
F_2	0.000100	0.000100	0.039572	0.000100
F_3	0.000020	0.000020	0.000020	0.000020
F_4	0.000014	0.000010	0.004469	0.000010
F_5	0.000250	0.000250	0.000250	0.000250
F_6	0.000700	0.000700	0.000700	0.000700
F_7	0.000035	0.006836	0.000035	0.000035
R_1	1.200000	0.781331	1.200000	1.200000
R_2	1.150000	0.089467	0.581567	1.150000
R_3	0.950000	0.950000	0.950000	0.950000
R_4	0.782086	0.424611	0.800000	0.800000
R_5	1.050000	0.531936	1.050000	1.050000
R_6	0.900000	0.900000	0.556429	0.900000
R_7	1.025000	0.591034	1.025000	1.025000

Table 3 Optimum value of availability using different metaheuristic algorithms on a swarm of 100 size with 1000 iterations

Algorithm	Population	Iterations	Availability	Elapsed time (s)
GWO	100	1000	0.998677	5.24
GOA	100	1000	0.9969397	33.56
DA	100	1000	0.9619688	7.84
WOA	100	1000	**0.9986771**	**0.83**

Table 4 Optimized values of failure and repair rates of the cooling tower subsystem using different swarm intelligence-based algorithms (N_Pop: 100, N_Iter: 1000)

Optimized value of failure and repair rates	GWO	GOA	DA	WOA
F_1	0.000200	0.000200	0.000200	0.000200
F_2	0.000100	0.000100	0.000100	0.000100
F_3	0.000020	0.000020	0.009500	0.000020
F_4	0.000122	0.000259	0.006282	0.000010
F_5	0.000250	0.000250	0.022694	0.000250
F_6	0.000700	0.000700	0.000700	0.000700
F_7	0.000035	0.001171	0.007000	0.000035
R_1	1.200000	1.200000	1.200000	1.200000
R_2	1.150000	1.039051	1.150000	1.150000
R_3	0.950000	0.950000	0.950000	0.950000
R_4	0.394100	0.028615	0.800000	0.800000
R_5	1.050000	1.050000	1.050000	1.050000
R_6	0.900000	0.528675	0.900000	0.900000
R_7	1.025000	1.025000	1.025000	1.025000

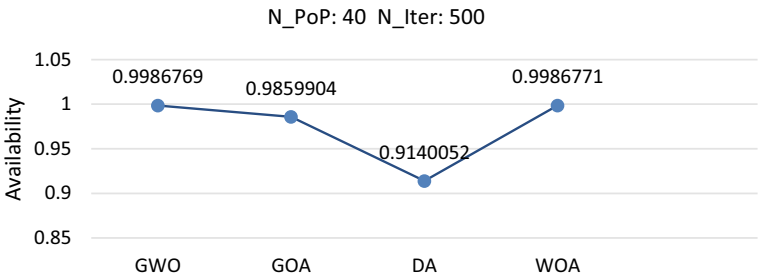


Fig. 1 Optimum value of availability using different metaheuristic algorithms with population size of 40 on 500 iterations

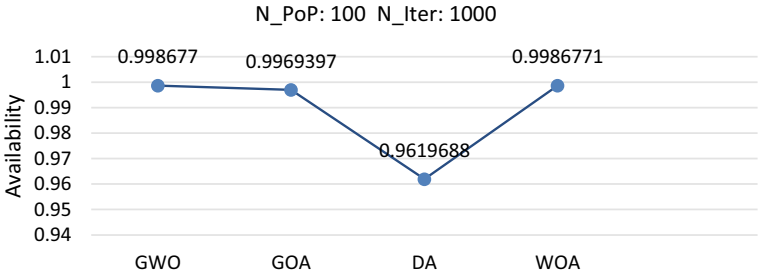


Fig. 2 Optimum value of availability using different metaheuristic algorithms with population size of 100 on 1000 iterations

4 Conclusion

Cooling towers play very critical role in the overall procedure of thermal power plants (TPP). To maximize the availability of TPP, it is vital that the failure and repair rate of various components of cooling tower should be optimized. Based on mathematical modelling and subsequent availability equations, the optimized values are obtained using four well-known SI-based metaheuristic approaches, viz. GWO, GOA, DA, and WOA. Experimental results shown that WOA outperformed other approaches in terms of maximizing the availability of cooling tower and time efficiency. Further, WOA can be utilized to maximize the availability of other process industries.

References

1. Khosravi F, Azli NA, Babaei E (2010) A new modeling method for reliability evaluation of Thermal Power Plants. In: 2010 IEEE international conference on power and energy, pp 555–560
2. de Souza GFM, Carazas FJG, dos S. Guimarães L, Rodriguez CEP (2012) Combined-cycle gas and steam turbine power plant reliability analysis. In: Thermal power plant performance analysis. Springer, pp 221–247
3. Saini M, Kumar A (2020) Stochastic modeling of a single-unit system operating under different environmental conditions subject to inspection and degradation. Proc Natl Acad Sci India Sect A Phys Sci 90(2):319–326
4. Gupta N, Saini M, Kumar A (2020) Operational availability analysis of generators in steam turbine power plants. SN Appl Sci 2(4):1–11
5. Gupta N, Kumar A, Saini M (2021) Reliability and maintainability investigation of generator in steam turbine power plant using RAMD analysis. J Phys Conf Ser 1714(1):12009
6. Khodakaram-Tafti A, Golneshan A-A (2020) A general mathematical model for predicting the thermal performance of natural draft dry cooling towers and extending it to three aligned towers. Energ Sour Part A Recover Util Environ Eff 1–23. https://doi.org/10.1080/15567036.2020.1844348.
7. Sharma A et al (2021) Opposition-based tunicate swarm algorithm for parameter optimization of solar cells. IEEE Access 9:125590–125602

8. Jagtap HP, Bewoor AK, Kumar R, Ahmadi MH, Chen L (2020) Performance analysis and availability optimization to improve maintenance schedule for the turbo-generator subsystem of a thermal power plant using particle swarm optimization. Reliab Eng Syst Saf 204:107130
9. Zhang H (2020) Transient performance of the particle swarm optimization algorithm from system dynamics point of view. J Comput Inf Sci Eng 20(4)
10. Saini M, Gupta N, Shankar VG, Kumar A Stochastic modeling and availability optimization of condenser used in steam turbine power plants using GA and PSO. Qual Reliab Eng Int. https://doi.org/10.1002/qre.3097
11. Kumar A et al (2022) Efficient stochastic model for operational availability optimization of cooling tower using metaheuristic algorithms. IEEE Access 10:24659–24677
12. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey Wolf optimizer. Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007
13. Saremi S, Mirjalili S, Lewis A (2017) Grasshopper optimisation algorithm: theory and application. Adv Eng Softw 105:30–47. https://doi.org/10.1016/j.advengsoft.2017.01.004
14. Mirjalili S (2016) Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Comput Appl 27(4):1053–1073. https://doi.org/10.1007/s00521-015-1920-1
15. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

Conditional Variational Autoencoder-Based Sampling

Firuz Kamalov, Adamu Ali-Gombe, and Sherif Moussa

Abstract Imbalanced data distribution implies an uneven distribution of class labels in data which can lead to classification bias in machine learning models. The present paper proposes an autoencoder-based sampling approach to balance the data. Concretely, the proposed method utilizes a conditional variational autoencoder (VAE) to learn the latent variables underpinning the distribution of minority labels. Then, the trained encoder is employed to produce new minority samples to equalize the sample distribution. The results of numerical experiments reveal the potency of the suggested technique on several datasets.

Keywords Imbalanced data · Sampling · Autoencoder

1 Introduction

Imbalanced data are a common problem in data science and machine learning applications. It refers to uneven distribution of target labels in data. Imbalanced data occur in various contexts including medical diagnostics, text classification, intrusion detection, image classification, and others. An uneven distribution of class labels in data can lead to classification bias in machine learning models [22]. Since the objective of a classifier is to minimize the overall classification error on the dataset, it will focus its learning on the majority labels to gain the maximum benefit. On the other hand, the minority labels are often of greater importance than the majority labels, and their classification requires higher priority.

F. Kamalov (✉) · S. Moussa
Department of Electrical Engineering, Canadian University Dubai, Dubai, UAE
e-mail: firuz@cud.ac.ae

S. Moussa
e-mail: smoussa@cud.ac.ae

A. Ali-Gombe
Mintra Research, Aberdeen, UK
e-mail: adamu.aligombe@mintra.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_66

There exists three main approaches to handle imbalanced data: (i) sampling, (ii) 1-class learning, and (iii) weighted cost function. While 1-class learning and weighted cost function approaches have been used in several studies [1, 15], the most popular approach in the literature is data sampling. Sampling refers to balancing the data either by the way of undersampling—reducing the number of majority samples—or via oversampling—increasing the number of the minority samples. Sampling techniques are an actively researched area with authors continuously proposing new algorithm in the field. Despite the large amount of effort dedicated to addressing this issue, it remains an open problem. In this study, we propose a sampling method based on conditional variational autoencoder (VAE). Numerical experiments demonstrate that the proposed technique improves classification accuracy on a number of datasets.

The majority of the existing sampling methods employ parametric models to generate new samples. As a result, the new sample points are distributed according to the model assumptions. However, the true underlying distribution of the data is rarely the same as the model distribution which skews the classification results. To deal with the issue of model bias, we suggest a data-driven approach to sampling. The proposed method relies on artificial intelligence—in the form of VAE—to learn the patterns within the data and generate new samples from the learned model. We believe that it allows for a more tailored approach to each dataset and consequently a better performance.

2 Literature

Sampling techniques can be grouped into two categories: (i) oversampling and (ii) undersampling [11]. Oversampling refers to leveling the data by artificially increasing the number of minority samples. The simplest approach to oversampling is by randomly choosing new points from the existing minority data (ROS). This approach, in effect, increases the weight of each minority point. The ROS method is a standard benchmark for measuring the efficacy of new sampling techniques. Another popular sampling benchmark is the SMOTE algorithm proposed in [3]. In SMOTE, the new points are randomly generated along a straight-line path between neighboring minority points. While the SMOTE algorithm implicitly assumes a uniform distribution of the minority points between the neighboring points, more advanced methods use other statistical distributions to produce the new minority samples [7, 10]. Alternatively, nonparametric techniques such as kernel density estimation can be used to estimate the spread of the minority samples [8, 9]. The authors in [20] suggest an approach based on weighting samples for imbalanced data in regression using KDE.

The use of generative models for sampling has been very limited. Most of the existing methods in this family of algorithms employ generative adversarial networks (GANs) to discern the underlying patterns within minority point distribution and use it to generate the new minority samples. Regular GAN was used for data augmentation in fault diagnosis methods on rotating machines [24]. A combination of GAN and Markov random fields was used in [19] to generate new minority points. As

an alternative to regular GANs and to address some of the related issues, conditional GAN was used for oversampling in [4]. Similarly, the conditional Wasserstein GAN was proposed and tested in [25]. Recently, the authors in [21] considered capsule GAN, where the network was trained using a loss function with embedded feature matching to accelerate training convergence. Their approach is shown to be particularly effective at recognizing highly overlapping classes using fewer parameters.

Given the large number for sampling techniques, several authors attempted to identify the optimal algorithm through comparative studies. In [13], the authors perform a large-scale comparison of SMOTE-based algorithms. A systemic evaluation of KDE-based sampling showed its robust performance compared to other approaches [18]. In [23], the authors compared different sampling approaches in the case of credit card fraud. The outcomes of comparative studies are inconclusive and depend on various factors. In this regard, the authors in [16] point to pitfalls of sampling.

3 Autoencoder-Based Sampling

3.1 Variational Autoencoder

Variational autoencoder (VAE) is designed to discern the intrinsic patterns within the data and generate new points based on the trained model [12]. As shown in Fig. 1, VAE contains two main parts: (i) encoder and (ii) decoder.

The encoder projects the variables in the original space onto a lower-dimensional latent space. By compressing the original variables into lower-dimensional features, the encoder aims to extract the important patterns hidden within the data. The encoder is often a multilayer perceptron (MLP) but can also be other type of a neural network such as convolutional (CNN) or recurrent (RNN). It is, in fact, a nonlinear generalization of principal component analysis (PCA).

The decoder is used to reconstruct samples in the original feature space based on the values from the lower-dimensional latent space. The decoder is usually designed as the inverse of the encoder—the layers of the decoder are in a symmetrically reverse

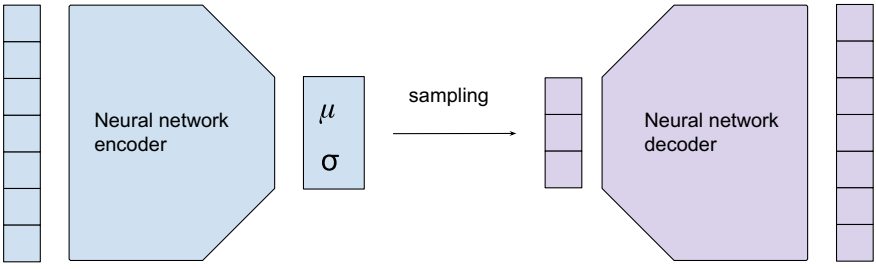


Fig. 1 The general VAE structure. The encoder maps each sample to a Gaussian distribution while the decoder samples from the latent distribution and attempts to reconstruct the original input

order than the encoder. While the encoder gradually reduces the dimensions of the hidden layers, the decoder gradually increases the dimensions of the hidden layers. Since the latent space is smooth, it can be used to sample new points. Thus, VAE is a generative neural network. VAEs have been used successfully in image and sound generation [6, 17]. Given their robust performance as generative models, VAEs are a natural candidate for minority sample generation.

Conditional VAE is based on the traditional VAE which facilitates the control of the data generation process. The conditional VAE employs a loss function akin to the original VAE:

$$L(x|c, \hat{x}) + \mathrm{KL}(q(z|x, c), p(z|c)), \tag{1}$$

where c is the conditional variable. In most cases, the data are conditioned on the labels. In other words, VAE is trained by taking into account the label information. It allows to subsequently generate new data according to the label. In particular, we can generate specifically minority labeled points for our sampling strategy.

3.2 VAE-Based Sampling

The proposed sampling algorithm consists of training a conditional VAE on imbalanced data and using the trained decoder to produce new minority observations to equalize the data. The structure of the proposed VAE is made up of the input layer, the bottleneck layer, and the output layer. Our numerical experiments show that even such an uncomplicated structure is capable of producing robust results. The VAE encoder used in our experiments consists of an MLP with a single hidden layer. As input, we concatenate features and one-hot encoding of the class label. The hidden layer is a dense ReLU-activated layer, while the output is a dense layer with no activation. The decoder is similar to the encoder. It receives and concatenates the encodings before passing them to the hidden layer which is a dense ReLU-activated layer, while the output layer is a dense sigmoid-activated layer. The VAE is trained using KL loss and a reconstruction loss.

After training the VAE on the train data, we use the trained decoder to generate new minority points. The number of newly produced minority samples is given by

$$n_{\text{new}} = N - n.$$

The steps of the proposed algorithm are summarized below:

1. Train VAE on the original imbalanced data.
2. Use the trained decoder to generate new minority samples.

4 Numerical Experiments

We evaluate the performance of the suggested sampling algorithm on a number of imbalanced datasets. The numerical experiments were carried out in Python using the TensorFlow library. The procedure for the numerical experiments is illustrated in Fig. 2. All the reported results are based on tenfold cross-validation.

4.1 Experimental Data

In our numerical experiments, we employ 4 different imbalanced datasets from a range of applications. All the datasets used in the experiments are publicly available in online repositories. The datasets have different characteristics in terms of the number of features, samples, and imbalance ratio. In addition, we employ both binary and multi-class data to obtain a more general experimental setup. The description of the datasets utilized in the study is provided in Table 1. For multi-class datasets, the imbalance ratio reported on the smallest class.

4.2 Benchmarking

To determine the effectiveness of the proposed sampling algorithm, we benchmark it against the original imbalanced data. In other words, we consider whether our sampling approach improves the classification accuracy over no sampling. In addition,

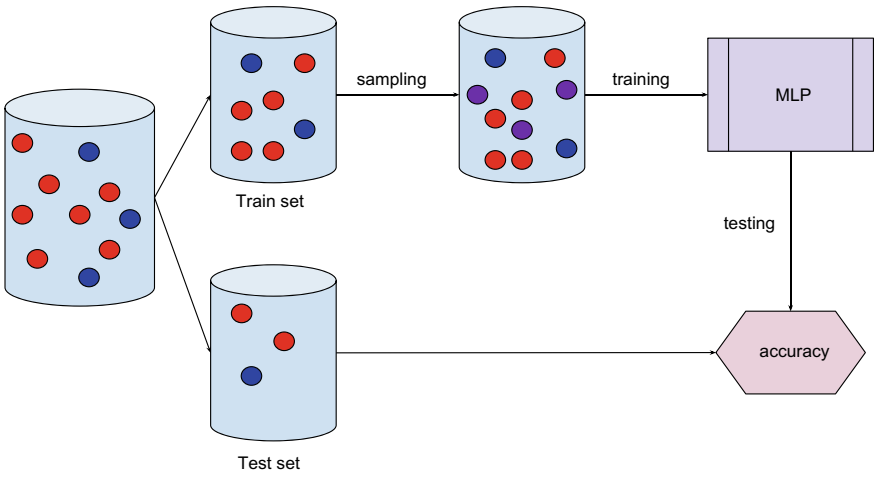


Fig. 2 The experimental procedure

Table 1 Details of the datasets used in the experiments

Dataset	No. samples	Imbalance ratio	No. features	No. classes	Source
Abalone	4024	1:12	11	14	UCI repo
Adult_census	48,842	1:3	13	2	UCI repo
Ecoli	336	1:72*	9	8	UCI repo
Mushroom	8124	1:42	22	2	UCI repo

we compare the performance of the suggested approach against a pair of popular methods: random oversampling (ROS) and random undersampling (RUS). ROS is a simple yet effective algorithm. It repeatedly selects new samples from the pool of the existing minority samples until a balanced data are reached. Thus, in ROS, a portion of the minority samples is represented several times. RUS is another simple yet effective sampling technique. It randomly selects samples the majority class. In RUS, the size of selected subset equals the number of minority samples in order to obtain a level set.

4.3 Results

We begin our discussion with binary datasets: *adult* and *mushroom*. The description of the datasets is provided in Table 1. We apply our sampling approach and train the MLP classifier on the balanced dataset as described in Fig. 2. The experiment is carried out via a tenfold cross-validation, and the mean accuracies are calculated. The results of the numerical experiments are presented in Table 2. The table shows that the VAE-based sampling method achieves the highest accuracy on both datasets: 0.8082 and 0.9877. The improvement in accuracy using the proposed method is statistically significant. The p-value for the accuracies on *adult* and *mushroom* is 8.8e−14 and 0.0024, respectively. The accuracy of the suggested approach is particularly greater than the ROS and RUS algorithms. On the other hand, the results based on other performance metrics are mixed. In both datasets, the proposed method has a relatively high precision rate but low recall rate. The results in Table 2 show that VAE-based sampling provides a robust alternative to the existing sampling algorithms.

Next, we consider multi-label datasets: *abalone* and *ecoli*. The details of the numerical experiments on multi-class data are presented in Table 3. The VAE sampling technique attains the highest accuracy with *abalone* and *ecoli* datasets. Concretely, the suggested approach has a substantially higher accuracy compared to the benchmark approaches on the *abalone* and *ecoli* datasets. The difference between the accuracy using the VAE sampling and the second best method is 0.05 and 0.21 on *abalone* and *ecoli* datasets, respectively. The improvement in accuracy using the proposed method is statistically significant with p-value less than e−16 in both

Table 2 The results of the numerical experiments on binary datasets

		Precision	Recall	f_1-score	Accuracy
Adult	Original	0.8558	0.5877	0.5910	0.8019
	Oversampling	0.7259	0.7961	0.7295	0.7644
	Undersampling	0.7101	0.7581	0.7156	0.7642
	VAE sampling	0.8492	0.6034	0.6156	**0.8082**
Mushroom	original	0.9930	0.7046	0.7748	0.9861
	Oversampling	0.4947	0.5414	0.4304	0.7596
	Undersampling	0.4947	0.5414	0.4304	0.7596
	VAE sampling	0.8852	0.7117	0.7647	**0.9877**

Table 3 The results of the numerical experiments on multi-class datasets

		Precision	Recall	f_1-score	Accuracy
Abalone	Original	0.0685	0.1087	0.0691	0.2047
	Oversampling	0.2372	0.2253	0.1839	0.2106
	Undersampling	0.0585	0.1142	0.0466	0.0889
	VAE	0.1617	0.1648	0.1414	**0.2533**
Ecoli	original	0.4132	0.4078	0.3420	0.5091
	Oversampling	0.1528	0.2503	0.1157	0.2212
	Undersampling	0.1528	0.2503	0.1157	0.2212
	VAE	0.7131	0.7118	0.6588	**0.7348**

cases. The precision and recall scores on the *abalone* dataset are the second highest. The best performance is shown on the *ecoli* dataset, where the proposed algorithm achieved the highest marks across all performance criteria.

The results demonstrate that the VAE-based sampling is able to improve the classification performance on imbalanced multi-class datasets by a significant margin. It is particularly effective in the case of structured data. However, the results are less impressive on unstructured data such as images. To improve the performance of VAE sampling on unstructured data a more sophisticated, multi-layered VAE is advised.

5 Conclusion

The present study considered a new sampling technique based on conditional variational autoencoder (VAE). In the proposed approach, a conditional VAE is trained on the imbalanced data. Then, the trained decoder is employed to produce new minority samples to level the data. The results of numerical experiments show that the

suggested approach attains strong performance. In particular, on the *abalone* and *ecoli*, the VAE sampling method produces significantly better performance than the benchmark approaches.

References

1. Aurelio YS, de Almeida GM, de Castro CL, Braga AP (2019) Learning from imbalanced data sets with weighted cross-entropy function. Neural Process Lett 50(2):1937–1949
2. Bagui S, Li K (2021) Resampling imbalanced data for network intrusion detection datasets. J Big Data 8(1):1–41
3. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority oversampling technique. J Artif Intell Res 16:321–357
4. Douzas G, Bacao F (2018) Effective data generation for imbalanced learning using conditional generative adversarial networks. Expert Syst Appl 91:464–471
5. Gan D, Shen J, An B, Xu M, Liu N (2020) Integrating TANBN with cost sensitive classification algorithm for imbalanced data in medical diagnosis. Comput Ind Eng 140:106266
6. Grekow J, Dimitrova-Grekow T (2021) Monophonic music generation with a given emotion using conditional variational autoencoder. IEEE Access 9:129088–129101
7. Kamalov F, Denisov D (2020) Gamma distribution-based sampling for imbalanced data. Knowl Based Syst 207:106368
8. Kamalov F (2020) Kernel density estimation based sampling for imbalanced class distribution. Inf Sci 512:1192–1201
9. Kamalov F, Elnagar A (2021) Kernel density estimation-based sampling for neural network classification. In: 2021 International symposium on networks, computers and communications (ISNCC). IEEE, pp 1–4
10. Kamalov F, Thabtah F, Leung HH (2022) Feature selection in imbalanced data. Ann Data Sci 1–15
11. Kaur H, Pannu HS, Malhi AK (2019) A systematic review on imbalanced data challenges in machine learning: applications and solutions. ACM Comput Surv (CSUR) 52(4):1–36
12. Kingma DP, Welling M (2013) Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114
13. Kovács G (2019) An empirical comparison and evaluation of minority oversampling techniques on a large number of imbalanced datasets. Appl Soft Comput 83:105662
14. Li Z, Huang M, Liu G, Jiang C (2021) A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection. Expert Syst Appl 175:114750
15. Lin WC, Tsai CF, Hu YH, Jhang JS (2017) Clustering-based undersampling in class-imbalanced data. Inf Sci 409:17–26
16. Moniz N, Monteiro H (2021) No free lunch in imbalanced learning. Knowl Based Syst 107222
17. Pu Y, Gan Z, Henao R, Yuan X, Li C, Stevens A, Carin L (2016) Variational autoencoder for deep learning of images, labels and captions. Adv Neural Inf Process Syst 29:2352–2360
18. Plesovskaya E, Ivanov S (2021) An empirical analysis of KDE-based generative models on small datasets. Procedia Comput Sci 193:442–452
19. Salazar A, Vergara L, Safont G (2021) Generative adversarial networks and Markov random fields for oversampling very small training sets. Expert Syst Appl 163:113819
20. Steininger M, Kobs K, Davidson P, Krause A, Hotho A (2021) Density-based weighting for imbalanced regression. Mach Learn 110(8):2187–2211
21. Shamsolmoali P, Zareapoor M, Shen L, Sadka AH, Yang J (2021) Imbalanced data learning by minority class augmentation using capsule adversarial networks. Neurocomputing 459:481–493

22. Thabtah F, Hammoud S, Kamalov F, Gonsalves A (2020) Data imbalance in classification: experimental evaluation. Inf Sci 513:429–441
23. Xiao J, Wang Y, Chen J, Xie L, Huang J (2021) Impact of resampling methods and classification models on the imbalanced credit scoring problems. Inf Sci 569:508–526
24. Zhang W, Li X, Jia XD, Ma H, Luo Z, Li X (2020) Machinery fault diagnosis with imbalanced data using deep generative adversarial networks. Measurement 152:107377
25. Zheng M, Li T, Zhu R, Tang Y, Tang M, Lin L, Ma Z (2020) Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification. Inf Sci 512:1009–1023

Discerning of In-Somnolence Using Body Sensors to Predict Vital Measurements

B. Prathap Kumar, P. Janani, K. Dhanush, Hanumantha Lakshmi Narayana, C. Karaka Teja, and K. Ganesh

Abstract In present situations, most people are suffering from insomnia. Due to stress, nowadays sleeplessness has a negative impact on human health. The medication treatment for this illness causes many side effects. Further, this leads to anxiety, depression, irritation, etc. It affects the performance of the user who is currently experiencing insomnia resulting in adverse conditions; it may even lead to accidents. This proposed work mainly discusses a non-pharmacological approach that reduces insomnia and provides better sleep which reduces side effects and implements better sleep. The latest diagnosing factors depicting the illness, even though it is lagging with certain conditioning factors, will be able to overcome those deviations from the existing technology. Therefore, this provides a betterment in the healthcare technology with the latest advancements for both the user and end-user applications. The results have provided the best results as it can be further developed as a wearable device in the future for the healthcare industry.

Keywords Insomnia · Arduino IDE · NodeMCU · Health care · Sleeplessness · Ubidots · IR sensor · Wearable body sensors

1 Introduction

Insomnia is the most common sleeping disorder in which it is difficult to fall and remain asleep. Grogginess in the morning is simply sleep inertia, which is a natural component of the walking process. In addition, the brain does not usually wake up immediately after sleeping. Insomniacs are more likely to suffer from heart disease and even excessive blood pressure [1]. It also raises the risk of having a stroke, which occurs when the brain's oxygen and blood supply is cut off owing to a clogged blood vessel. Anxiety about finances, health, work or education will keep the brain active at night, and it makes it hard to sleep. Insomnia can also be caused by a stressful

B. Prathap Kumar (✉) · P. Janani · K. Dhanush · H. Lakshmi Narayana · C. Karaka Teja · K. Ganesh
Vel Tech Rangarajan Dr Sagunthala R and D Institute of Science and Technology, Chennai, Tamil Nadu, India
e-mail: bprathapkumar66@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_67

life, such as illness of loved ones, job loss or financial loss also which may cause to sleeplessness. Poor sleep habits, such as an inconsistent bedtime routine, naps before bed and an unpleasant sleep environment, such as sleep at working before bed, using cell phones or laptops before night, watching TV or playing video games before bed can all cause sleep disruption. Insomnia causes breathing problems all over the night [2]. Unpleasant sensations in the legs are caused by "Restless legs syndrome", and it is very difficult to move them, to prevent falling asleep. And avoiding some actions like stop drinking alcohol at nights helps to sleep, but it causes awakening in the middle of night, avoid preferring heavy foods at night, avoid drinking coffee, tea, caffeinated drinks at late nights [3].

Some things to do to reduce insomnia include doing some body exercises, sticking to a regular sleep schedule, getting plenty of sunlight during early morning, taking a warm bath before bed, drinking warm milk and listening to music, meditating or exercising which can all help to alleviate insomnia in a natural manner [4]. Insomnia is a prevalent health concern in both the medical and psychiatric fields. Insomnia is defined as a discontent with the amount or quality of sleep, as well as trouble falling asleep, staying asleep and getting up in the morning [5]. It has always been undervalued and trivialized in terms of its health consequences. Moreover, insomnia could cause a decrease in daytime function and a significant mental and physical disorders for the patients. Electronic sensors are rapidly improving to allow IoT in health telemonitoring, which should meet the above standards. The application of the anytime, anyplace idea could provide patients with insomnia with the freedom and expression of clinical and nonclinical data on a continual basis [6]. Body sensor networks are a new technology that allows for long-term biosensor monitoring by not interfering with a person's routine works [7]. It combines various sensors with the human body to provide a complete solution for providing relevant information about personalized health, tracking personal well-being, detecting health risks and facilitating rapid response in severe events such as stroke, epilepsy and heart attack and so on [8].

2 Problem Statement

Chronic insomnia is determined as insomnia that takes place at least three nights per week in a month or more. Sleeplessness can be caused by a variety of circumstances, including life stress (like job loss or change, death of loved ones, etc.), illness and environmental factors such as noise, light or very high temperatures. Long-term insomnia can be caused by chronic stress and depression [9]. A conditioned emotional response is a typical cause of insomnia. Thoughts about sleeping difficulties and the behaviours that arise are given as a result of them [10]. Insomnia can be caused by a slow rate of reaction when driving, which increases the danger of an accident. As a result, insomnia can occur at any moment during the cancer phase, included during the initial diagnosis, surgery, treatment, during diagnosis or treatment of a recurrence and terminal stages of the disease [11].

3 Insomonolence

The author has discussed the development of system for analysis of patients with insomnia. The system is composed with multi-body sensors that are connected to the IOT server. The server is enforced by using representational state transfer (REST) Api for robust and lightweight data transferring from the device. Arulvallal et al. [12] authors have proposed a continual sleep monitoring system that detects spo2 and blood pressure and heart rate. These parameters could be displayed by using a liquid crystal screen and could be wirelessly sent from the system to the smartphone for additional examination via Bluetooth or Wi-Fi. Lokavee et al. [13] introduced system-supported pressure sensors that are embedded inside a pillow sheet which will provide a not noticeable method for monitoring sleep. The system consists of wireless network devices, pressure sensors and software to research and classify bodily movements. Li et al. proposed a bed-mounted seismometer device for tracking heart and breathing rates, as well as body movement and posture, while sleeping [14]. This system is non-contact and non-intrusive, demonstrating strong potentials in terms of sleep quality and monitoring. It also detects a user's minor activities such as snoring while sleeping. Parlikar [15] presented a practical strategy for dealing with insomnia and ensuring that the user gets a healthy sleep, and also, a relaxation therapy is used to overcome insomnia, resulting in increased quality of sleep. Hsieh has introduced a sleep quality detection system that takes out the user's sleep information and sends to the cloud platform [16]. This system consists of Internet of Things (IoT), and the cloud platform and sensors are connected to device [17].

Subsequently, medical staff and user can access the information of quality of their sleep [18]. Suciu et al. [19] discussed three solutions for monitoring the quality of life, and they are Adcon, Fitbit and uRADMonitor. Moyosola et al. [20] have implemented an user-friendly neurofeedback tool which is capable of treating mental and physical issues like physical pain and mental issues like addiction to a certain drug, over aggression, anxiety due to various factors, autism, depression, hyperactivity, insomnia and can also be in case of brain damage caused by EEG signals. This analysis and detecting the presence of EEG can be done through home-made neurofeedback system [21].

4 Proposed Methodology

Insomnia is defined as a discontent with the amount or quality of sleep, as well as trouble falling asleep, staying asleep and getting up in the morning. This system's design was split into two parts: hardware and software. The first step was to choose adequate sensors. The user's eye movement is detected using an eye blink sensor [22]. This eye blink sensor is based on infrared technology. As the eye blinks, the variation throughout the eye will change [23]. If the eye is closed, the output is high; otherwise, the output will be low. This determines whether the eye is shutting or

opening. A buzzer, also known as beeper, is an audio signalling device which can be electromechanical, electronic or mechanical. Before Arduino, microcontrollers did not have a software IDE for uploading code to the hardware [24]. To upload the code into the hardware, one needed to use a separate hardware device. It is simple to connect sensors to Arduino because of this flexibility feature. When it comes to creating a new project, Arduino is the first choice for students and adults who are new to electronics. Other microcontrollers can also be utilized with Arduino sensors. The Arduino IDE includes several libraries that can be used to connect various sorts of sensors. The sensors that require high processing speed than the Arduino can give are the sole exceptions (Fig. 1).

The Arduino IDE software and the embedded C programming language make up the second section. The zero PCB board is given hardware connections. NodeMCU is connected to a Wi-Fi hotspot. Through NodeMCU, a USB cable is linked to the PC. The user becomes drowsy as a result of a severe workload, which can be avoided by using an eye blink sensor [25]. The software used is called Arduino IDE. It is a text editor with a variety of capabilities, like a notepad. It is used to write programmes, compile them to check for errors and then upload the code to the Arduino. The Arduino IDE Software executes embedded C code, which displays the user's sleep stage as an output while also displaying user data in Ubidots. The system detects how many eye blinks are detected using a sensor and sends that information to the cloud platform Ubidots, from which we send an SMS to the higher authorities [26] (Fig. 2).

The main purpose of the developed system is to provide a sleep monitoring system which is based on loT. This is a system which works on cloud and is created to enhance the features which are currently unavailable. The system is also capable to monitor the variation in the range of the sensors. The system is categorized into two divisions [27].

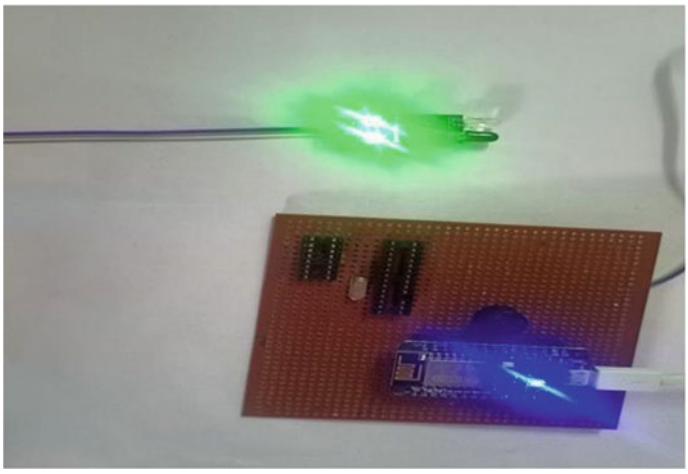


Fig. 1 Hardware setup using Arduino

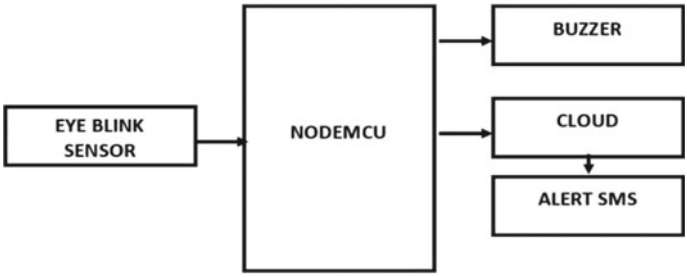


Fig. 2 Block diagram of system

- Sensor subsystem
- Web service

4.1 Sensor Subsystem

The sensor network contains a sensor: IR sensor module (infrared sensor). The IR sensor detects motion as well as measuring the heat of objects. It contains an infrared LED and an infrared photodiode, which can be combined to make a photocoupler or an optocoupler. The IR sensor module includes five essential parts like IR, Tx, Rx, operational amplifier, trimmer pot (variable resistor) and output LED. The pin configuration of the IR sensor module is shown below. (a) VCC pin is input power supply; (b) GND Pin is ground power supply; (c) OUT is an active-high o/p [28].

4.2 Web Service

Wi-Fi module is used to communicate between NodeMCU and Arduino IDE over a network. Ubidots is the cloud platform used. A link will be made available to each user to access the Website dashboard. As it is web based and is in cloud, it is available through all kinds of devices like laptops and mobiles. The IR sensor and buzzer are connected to the NodeMCU controller. IR sensor has three pins; in that GND and VCC are the power supply, and one data pin is connected to the NodeMCU digital pin. And, buzzer has 2 pins. One is GND, and another one is digital pin of NodeMCU. A Universal Serial Bus (USB) cable is connected to Node MCU and system (PC or laptop installed with Arduino IDE software). An embedded C programming code is given in Arduino IDE. After execution of code in Arduino IDE, the IR sensor detects the user eye blink movement and a green colour light glows on IR sensor. Also, the ESP2866 gets activated and a blue colour light is seen on it. Now connect a Wi-Fi hotspot to ESP2866 [29].

Fig. 3 Insomnolence alert message

The buzzer makes a loud audio signal alert, and the hospital staff gets notified. Now open Ubidots cloud platform and create your account and log in on the Website. Now click on devices and again select devices among all the options. Create a new device and variable using the following: Create device and click on blank device; name your device as "Smart home" then click on add variable; select raw and name your variable as "Bulb"; click on data and go to dashboards and create a widget using the following procedure: click on add new widget and select switch; name your widget as "Bulb" then click on add variables; select smart home and select "Bulb"; and then click on green tick mark. Then, select send Email option and enter an Email ID and click next. An insomnolence alert message will be sent to the given Email ID [30] (Fig. 3).

5 Conclusion

The system can detect whether a person is having good sleep or not by using IR sensor and NodeMCU. Sleep is important not only to physical but also to mental well-being. A good sleep is beneficial to all types of learning since it aids in attention, decision-making and creativity [31].

6 Result

The expected results of discerning of insomnolence using body sensors which are enabled to predict vital measurements should be defined from IR sensor which predicts the eye blink movement of the user. It is important to keep track of your sleep status and quality in order to better understand your health and to prevent sleep disorders. It is necessary to develop health monitoring sensors that are comfortable, easy to use and inconspicuous. Wearable sensors are used to monitor vital bio-signals. The system determines whether the user is having healthy sleep or not with the help of IR sensor and NodeMCU (Figs. 4 and 5).

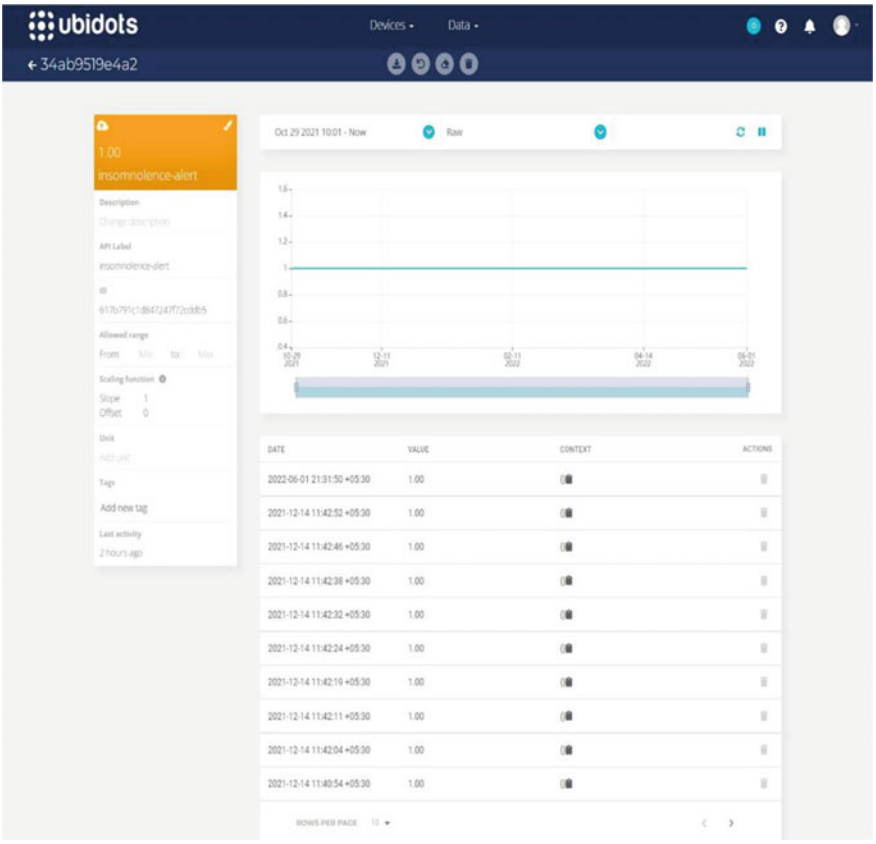


Fig. 4 Output graph in Ubidot

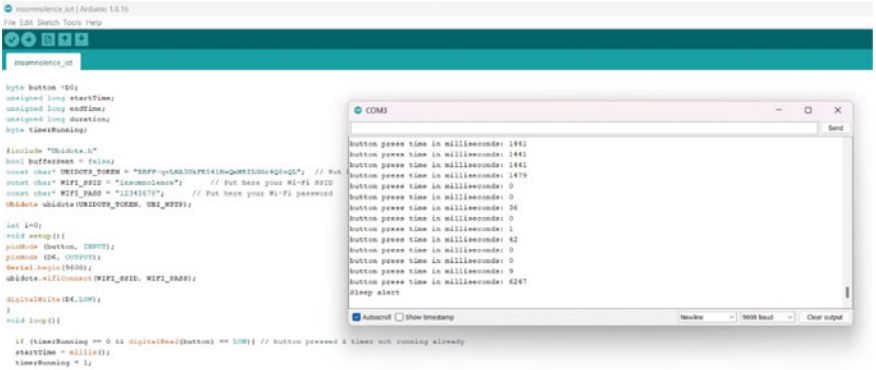


Fig. 5 Arduino IDE output

References

1. Li S, Chiu C (2018) A smart pillow for health sensing system based on temperature and humidity sensors. Res Gate Sens 18(11):3664
2. Deep S, Zheng X, Karmakar C, Yu D, Hamey LG, Jin J (2019) A survey on anomalous behavior detection for elederly care using dense-sensing networks. 22(1), Firstquarter 2020, IEEE Commun Surv Tutorials Author, F.: Contribution title. In: 9th international proceedings on proceedings, pp 1–2
3. Lauteslager T, Kampakis S, Williams AJ, Maslik M, Siddiqui F (2020) FRCP FAASM. In: 2020 IEEE 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)
4. Veiga A, Garcia L, Parra L, Lloret J, Augele V (2018) An IoT-based smart pillow for sleep quality monitoring in AAL environments. In: 2018 IEEE third international conference on fog and mobile edge computing (FMEC)
5. El-Khadiri Y, Corona G, Rose C, Charpillet F (2018) Sleep activity recognition using binary motion sensors. IEEE 30th international conference on tools with artificial intelligence (ICTAI)
6. Zhang Y, Yang Z, Lan K, Liu X, Zhang Z, Li P, Cao D, Zheng J, Pan J (2019) Sleep stage classification using bidirectional LSTM in wearable multi-sensor systems. In: IEEE INFOCOM 2019—IEEE conference on computer communications workshops (INFOCOM WKSHPS).
7. Kajornkasirat S, Chanapai N, Hnusuwan B (2018) Smart health monitoring system with IoT. In: IEEE symposium on computer applications & industrial electronics (ISCAIE)
8. Liu W, Chang S, Liu Y, Zhang H (2020) WiPSG: detecting rhythmic movement disorder using COTS Wi-Fi. IEEE Internet Things J 8(6)
9. Rahman MJ, Morshed BI (2021) A novel method for sleep score estimation using wearable sensors with a deep sequential neural network. In: 2021 IEEE international conference on electro information technology (EIT)
10. Schneider J, Schroth M, Ottenbacher J, Stork W (2018) A novel wearable sensor device for continuous monitoring of cardiac activity during sleep. In: 2018 IEEE sensors applications symposium (SAS)
11. Zhang F, Wu C, Wang B, Wu M, Bugos D, Zhang H, Liu KR (2019) SMARS: sleep monitoring via ambient radio signals. 20(1) IEEE Trans Mobile Comput
12. Arulvallal S, Snekhalatha U, Rajalakshmi T (2019) Design and development of wearable device for continuous monitoring of sleep APNEA disorder. In: 2019 IEEE international conference and signal processing (ICCSP)

13. Lokavee S, Tantrakul V, Suwansathit W, Pengjiam J, Kerdcharoen T (2019) Detection of sleep-disordered breathing and sleep efficiency with wireless pillow-sheet sensor network. In: 2019 IEEE 16th international conference on electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON)
14. Li F, Clemente J, Song W (2018) Non-intrusive and non-contact sleep monitoring with sesimometer. In: 2018 IEEE global conference on signal and information processing (GlobalSIP)
15. Parlikar S (2020) Occipital osteopathy and radiation detection for reducing insomnolence using IoT enabled pillow. In: 2020 IEEE fourth international conference on ISMAC (IoT in Social, Mobile, Analytics and Cloud) (ISMAC)
16. Hsieh YZ (2018) Internet of Things pillow detecting sleeping quality. In: IEEE 2018 1st international cognitive cities conference (IC3)
17. Chen C, Wang Z, Li W, Chen H, Mei Z, Yuan W, Tao L, Zhao Y, Huang G, Mei Y, Cao Z, Wang R, Chen W (2018) Novel flexible material-based unobtrusive and wearable body sensor networks for vital sign monitoring. IEEE Sens J 19(19)
18. Azman N, Satria MH, Mukaram MZ, Abd Ghani MK (2016) Development of embedded system for centralized insomnia system 2018. vol 9999. IEEE Springer, Heidelberg, pp 1–13
19. Suciu G, Petre I, Pasat A, Vulpe A, Obreja S (2018) Retrieving health and environmental data using secure IoT cloud platform. In: 2018 IEEE international symposium on electronics and telecommunications (ISETC)
20. Moyosola SM, Alexandru MD, Nicole G, Al Gayar Sarmad Monadel S, Pavaloiu B, Boiangiu C (2019) Development of a low-cost and user friendly neurofeedback tool to treat depression, insomnia, anxiety, pain and ADHD using an arduino and android application".
21. Yi R, Enayati M, Keller JM, Popescu M, Skubic M (2019) Non-invasive in-home sleep stage classification using a ballistocardiography bed sensor. In: 2019 IEEE EMBS international conference on biomedical & health information (BHI)
22. Ueafuea K, Boonnag C, Sudhawiyangkul T, Leelaarporn P, Gulistan A, Chen W, Mukhopadhyay SC, Wilaiprasitporn T, Piyayotai S (2020) Potential applications of mobile and wearable devices for psychological support during the COVID-19 pandemic. IEEE Sens J 21(6):7162–7178
23. Malik A, Naqvi SZ, Aziz S, Khan MU, Hussain T, Hasan F (2021) Detection of insomnia using electrocardiography and electromyography. In: IEEE 2021 international conference on artificial intelligence (ICAI)
24. Abdullah H, Penzel T, Cvetkovic D (2018) Time and frequency analysis of heart rate variability at different sleep stages of primary insomnia. In: IEEE 2018 2nd international conference on biosignal analysis, processing and systems (ICBAPS)
25. Raju NP, Venkatesh U, Yadhav S (2019) Diagnosing insomnia using single channel EEG signal. In: IEEE 2019 international conference on communication and electronics systems (ICCES)
26. Yang B, Liu H (2020) Automatic identification of insomnia based on single-channel EEG labelled with sleep stage annotations. IEEE Access 8
27. Kuo CE, Chen GT (2020) A short-time insomnia detection system based on sleep EOG with RCMSE analysis. IEEE Access 8
28. Mamta P, Prasad SV (2020) EEG feature extraction using time domain analysis for classifying insomnia. In: 2020 IEEE India council international subsections conference (INDISCON)
29. Vardhan VV, Venkatesh U, Yadav S (2020) Signal processing based autonomous sensory meridian response to treat insomnia. In: IEEE international conference on electronics and sustainable communication systems (ICESC)
30. Mahanta D, Bordoloi H, Saikia SJ (2020) LabVIEW based smart pillow. In: 2020 IEEE international conference on computational performance evaluation (ComPE)
31. Shahin M, Mulaffer L, Penzel T, Ahmed B (2018) A two stage approach for the automatic detection of insomnia. In: IEEE 40th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)

PROPHETESS: A Tool for Prediction of Prophage Loci in Bacterial Genomes

Manu Rajan Nair and T. Amudha

Abstract This paper describes the design, development, and implementation of a standalone bioinformatic tool for the prediction of putative prophage loci in bacterial host genomes using statistical measures, based on the algorithm published as the "Prophage Loci Predictor for Bacterial Genomes" and described as the loci predictor algorithm. This algorithm proposed a novel approach to the problem of detecting prophage regions in bacterial genomic information using particle swarm optimization, using a fixed size pattern lookup table to detect virus-like pattern distributions in the host/bacterial genome. As this algorithm was designed with the intension of providing highly consistent and fast performance, the time-to-process sequence is the primary metric for evaluating the performance of the tool, and the processing speed was expected to scale only with the size of the genome under consideration and not on the size of the pattern database as is the case with other algorithms in its class. The implemented tool was evaluated using both the algorithms test and training sets and was shown to obtain a linear co-related performance as expected in both training and prediction phases of the performance testing.

Keywords Bioinformatics · Genomics · Prophages · PSO

1 Introduction

The PROPHETESS tool is an implementation of the loci predictor algorithm (LPA) [1], and this paper outlines the various steps such as the design, development, and the implementation of this algorithm as a standalone tool for bioinformatic researchers. The primary aim of the LPA is to predict nucleotide loci within an un-annotated bacterial genome that contains high probability of containing virus remnants known as prophages [2]. The study of prophages and the complex interconnection between host, phage, and prophage has been a major topic of research over several fields such as virology, genomics, and bioinformatics. Prophages represent the intersection

M. R. Nair (✉) · T. Amudha
Department of Computer Applications, Bharathiar University, Coimbatore 641 046, India
e-mail: manurajan10@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_68

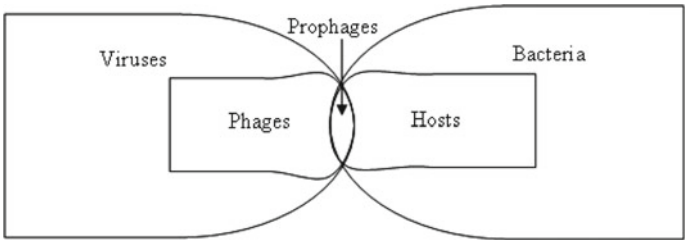


Fig. 1 Visualization of the distribution of information modules between the domains

of the bioinformatic information domain of bacteria and viruses, as they contain informational modules that may be shared by both species [3]. The extend of the overlap between these two sets is still one of the biggest fields of study in comparative genomics, consequently facing the information overload challenges that apply to this field. Figure 1 outlays the distribution of information.

Prophages share a common set of information modules with both bacteria (hosts) and their bacteriophages (phages), therefore, differentiating information between them becomes difficult taking into account the fact that the known does not represent the entire corpus of bioinformatic information represented by all three, the number of known viruses and bacteria, the known phages and their hosts, the known prophages; all together represents only a small percentage of all that exists in the biosphere. It has been shown that prophages play a significant part in the evolution of both viruses and bacteria [4], thus providing a shortcut for bioinformatic researchers to study both bacteria and viruses, their interaction, and evolution by studying these prophages.

The earlier methods for studying prophages involved direct laboratory experimentation and were obtained by processing bacterial cultures. These methods were inefficient in identifying all the prophages in a genome and were later replaced with more efficient methods as they became available to genomics. The development of bioinformatic tools and their application in genomics along with the development of automatic rapid sequencing techniques revolutionized the study of prophages. Now, prophages can be studied by the application of bioinformatic algorithms and tools to genomic information in electronic formats enabling quick identification and annotation of metagenomics.

These advances in bioinformatics have led to the development of a host of tools to help process and annotate information, starting with alignment tools used in automated sequencing machines to produce complete sequences of a genome from a biological sample to protein shape prediction at the other end of the pipeline. A specific set of tools were developed for identifying prophages in bacterial genomes, among them are Phage_Finder [5], PhiSpy [6], and ProphET [7]. Most of these tools use metagenomic information such as protein sequences as identifiers along with structural peculiarities such as flanking attachment sites as the primary source for identifying prophages. This means that they are limited in their ability to identify novel prophages, and their reliability is directly tied to the availability of metagenomic information. Another interesting feature about these tools is that they often

rule out all regions that do not have identifiable protein sequences to greatly reduce the search space, thereby increasing the speed of the search but at the cost of disregarding regions that contain degraded prophage information. It can also be observed that these tools are all suites that utilize pipelines of external general purpose tools such as BLAST [8] to do a lot of the heavy lifting, particularly when it comes to processing protein alignments and often serving only lighter algorithms to process the results the external tools in the pipeline. PHAST and its later versions such as PHASTER [9] are available as publicly accessible online web applications for the processing of prophage information, whereas tools like PhiSpy are classified as standalone, in the sense that they are installed on desktop class machines. Several purely statistical methods have been developed such as that using dinucleotide A+T/G+C abundance as identifiers [10], but such methods are intrinsically slow in terms of real-world performance. This shows that there is the need to develop a new algorithm that can process entire raw bacterial genomes and produce reliable statistical output on putative prophage areas or loci, and a tool that can be used as part of a pipeline to annotate prophages. The LPA was developed considering the limitations of the existing systems for prophage detection, and the aim of this research is to implement this algorithm as the PROPHETESS tool.

2 Methodology

The PROPHETESS tool is implemented as a set of modules containing core LPA modules and additional modules for functionalities such as evaluation. The nucleotide sequence hashing (NSH) is an LPA technique that converts the input FASTA files containing nucleotide sequences into a series of pre-selected ideal sequence encoding (ISE) codes. The various modules are as follows:

2.1 Preprocessing or Pre-preparation Module (Core LPA)

In this module, the input genomes, both bacterial host and viral phage, are processed in NSH files. Other miscellaneous preparatory checks are also done here, from this point on in the tool, genomes are handles as their NSH file only.

2.2 Training Module (Core LPA)

This module takes as input the training lists of host and phage NSH files and outputs the NSH pattern table which acts as a type of pattern database encapsulating the patterns learned by processing the training set data. This is a one-time operation for a given corpus of research in the sense that that if the training set or the ISE codes

are not deliberately changed, there is no need to repeat this step. All further research done are valid for this input training set.

2.3 Prediction Module (Core LPA)

This module is the operation end of the tool, it takes as input a testing set of host genomes, practically this consists of genomes that are novel and commits two levels of predictions on it, the first is the raw prediction where a large set of possible sites that may contain prophage data is predicted, in the second level, a particle swarm optimization (PSO)-based optimizer operates on this raw prediction and outputs the tentative final prediction. An overview of the overall design is shown in Fig. 2.

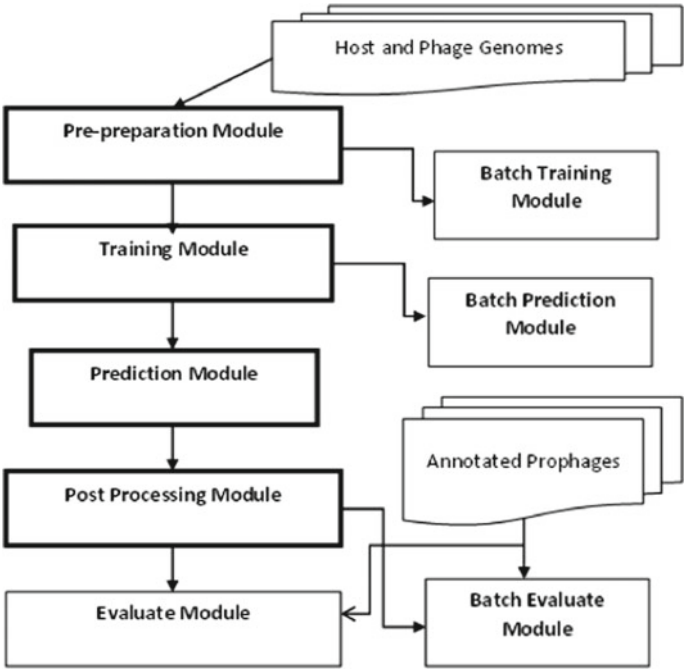


Fig. 2 PROPHETESS modules

2.4 Post-processing Module (Core LPA)

This module has the duty of consolidating the tentative final prediction into a proper final loci prediction by decoding the NSH positions into proper FASTA positions in the original input genome file.

2.5 Batch Training Module

This module provides the ability to process a training set given a CSV list of hosts and phages. This is for batch processing of training data.

2.6 Batch Prediction Module

This module provides the ability to process a testing set as whole given a CSV list of phages. This is for batch processing of testing data.

2.7 Evaluation Module

This module is used for evaluation or comparative study, and it produces a text output after calculating standard statistical measures such as false positives, false negatives, and true positives and calculates metric such as sensitivity and positive predictive value (PPV)/accuracy of the test set results.

2.8 Batch Evaluation Module

This module provides the ability to produce a set of results for a given CSV list of test set hosts.

3 Results and Discussion

The PROPHETESS tool was tested, and the time taken for each phase of operation was calculated. The success metric for this paper was based on the ability of the implementation to maintain performance without significant degradation of

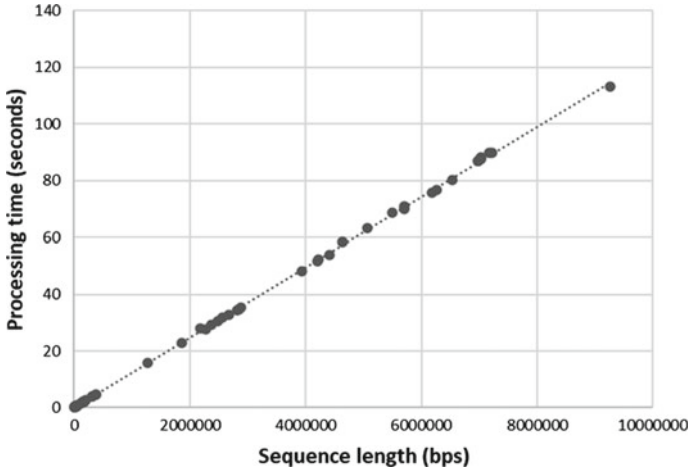


Fig. 3 Time-to-process training sequence versus sequence length

processing times. The four major time taken calculations were then analyzed against their respective data sequence size in bps, the metric chosen are the following:

1. Time-to-process Training sequence

This is defined as the time in seconds taken by the training module to finish processing each item in the training set, no distinction is made between host or phage in this metric. The graph is shown in Fig. 3. Each element represents a host or phage, the FASTA sequence length of that element on the X-axis and the respective processing time on the Y-axis.

2. Time-to-generate NSH

This is defined as the time taken by the training module to convert a preprocessed input file into an equivalent NSH. The graph is shown in Fig. 4.

3. Time-to-generate Raw Prediction

This is defined as the time taken by the prediction module to generate a raw prediction as shown is Fig. 5.

4. Time-to-generate Final Prediction

This is defined as the time taken to produce a fine-tuned output from a raw prediction by the prediction module. The graph is shown in Fig. 6.

An analysis of the above graphs reveals that the trend line for Fig. 3 is correlated with the data points suggesting that the implementation of the tool is consistent with the LPA expectations, the generation time is consistent with the size of the input data, the size of the input data, in this case, the sequence lengths of each member

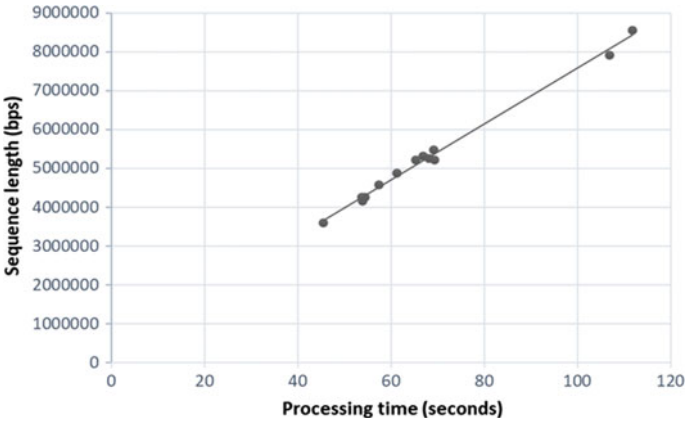


Fig. 4 Time-to-generate NSH versus sequence length

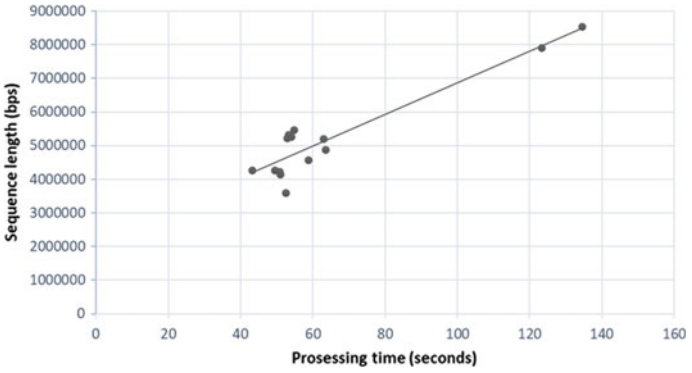


Fig. 5 Time-to-generate raw prediction versus sequence length

of the LPA training set are the only discriminant in determining the processing time and hence the performance of the tool.

The time-to-generate NSH graph represented by Fig. 4 also has a consistent trend line, and the same conclusion that the size of the input sequence is the only determinant in the performance of this phase of LPA. The time-to-generate raw prediction graph in Fig. 5 also shows clustering around the trend line, with only one (Ascension ID: CP030239) is the only one that seems a bit off. This is a Paracoccus bacteria and the smallest of the Blind Testing set from the LPA dataset.

The time-to-generate final prediction graph represented by Fig. 6 shows significant distress, particularly when handling the blind test set from the LPA dataset, but in this case, the two biggest outliers are also the two biggest data sequences under consideration, the anomalous nature can be attributed to the PSO implementation in the tool and also the fundamental non-stochastic nature of bio-inspired algorithms

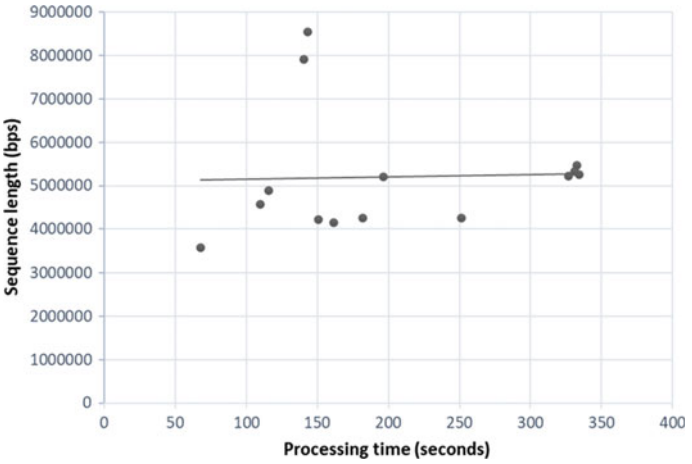


Fig. 6 Time-to-generate final prediction versus sequence length

in general. Overall, it can be assessed that the tool performed consistently while handling the training set and blind testing set as published in the LPA.

4 Conclusion

In conclusion, the PROPHETESS implementation closely follows the design paradigm set by the LPA algorithm, producing a tool with a few additional features that will allow it to serve as frontline tool for bioinformatic researchers. The primary thrust of the LPA was the faster speed of the algorithm in comparison to other standalone tools of its class, allowing researchers to go for a more iterative approach to the process of locating prophages in novel un-annotated bacterial genomes.

References

1. Nair MR, Amudha T (2021) Prophage loci predictor for bacterial genomes. J Bioinform Comput Biol 19(2):2050049 Epub 27 Feb 2021. PMID: 33641637. https://doi.org/10.1142/S0219720020500493.
2. D'Hérelles F (1949) The bacteriophage (PDF) Sci News 14:44–59. Retrieved 5 Sept 2010
3. Menouni R, Hutinet G, Petit MA, Ansaldi M (2015) Bacterial genome remodelling through bacteriophage recombination. FEMS Microbiol Lett 362(1):1–10. https://doi.org/10.1093/femsle/fnu022
4. Cole ST, Saint Girons I (1994) Bacterial genomics. FEMS Microbiol Rev 14(2):139–160. https://doi.org/10.1111/j.1574-6976.1994.tb00084.x PMID: 8049097
5. Casjens S (2003) Prophages and bacterial genomics: what have we learned so far? Mol Microbiol 49:277–300

6. Reis-Cunha JL, Bartholomeu DC, Manson AL, Earl AM, Cerqueira GC (2019) ProphET, prophage estimation tool: a stand-alone prophage sequence prediction tool with self-updating reference database. PLoS ONE 14(10):e0223364
7. Lima-Mendez G, Van Helden J, Toussaint A, Leplae R (2008) Prophinder, a computational tool for prophage prediction in prokaryotic genomes. Bioinformatics 24:863–865
8. Fouts DE (2006) Phage_Finder: automated identification and classification of prophage regions in complete bacterial genome sequences. Nucl Acids Res 34(20):5839–5851 Epub 24 Oct 2006, PMID: 17062630; PMCID: PMC1635311. https://doi.org/10.1093/nar/gkl732
9. Zhou Y, Liang Y, Lynch K, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. Nucl Acids Res 39(suppl 2):W347–W352. https://doi.org/10.1093/nar/gkr485 PMID: 21672955
10. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

IOT-Based Third Eye Glove for Smart Monitoring

S. Sudharsan, M. Arulmozhi, C. Amutha, and R. Sathya

Abstract The Internet of Things (IoT)-based third eye glove is the solution for the blind and employs ultrasonic sensors and an Arduino UNO board. The goal of the IoT is to link physical items, such as computers, to other physical objects, such as other people. Currently, it is a market-enabling technology that is quickly evolving and increasing. As many as 1.6 million youngsters are amongst India's 40 million blind people. Blind people have a tough time travelling on their own. They have to rely on others in many aspects of their existence. When they're strolling along the street, this is the biggest problem. With a stick in hand, they will not be able to see every obstruction. One may wear this sophisticated glove design for a long time. The third eye glove will assist the blind individual in achieving their goal. Sensors, microcontrollers, and buzzers are all integrated into the glove's design, which is IoT. The sneaker emits a buzzing sound as the user approaches an obstruction. Smart gloves, which have sensors and can cover a larger area, are being developed to increase productivity. Smart health monitoring and smart gloves ensure that the user does not have any problems.

Keywords Blind and deaf · IoT · Sensor · Glove

S. Sudharsan (✉) · M. Arulmozhi · R. Sathya
Department of ECE, Rajalakshmi College of Engineering, Rajalakshmi Nager, Thandalam, Chennai, Tamil Nadu, India
e-mail: sudharsan.s@rajalakshmi.edu.in

M. Arulmozhi
e-mail: arulmozhi.m@rajalakshmi.edu.in

R. Sathya
e-mail: sathya.r@rajalakshmi.edu.in

C. Amutha
Department of EEE, Rajalakshmi College of Engineering, Rajalakshmi Nager, Thandalam, Chennai, Tamil Nadu, India
e-mail: amutha.c@rajalakshmi.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_69

1 Introduction

The world's greatest population of blind persons today resides in India. India accounts for about 15 million of the world's 37 million blind people. For visually impaired persons to live independently, assistive technologies are being created. Blind persons can live normal lives and do things according to their preferences, but they confront many challenges compared to others who are not disabled. One of the most significant difficulties for visually impaired people, particularly those who are completely blind, is the inability to operate a smartphone. The IoT-based third eye glove aims to create a product that will be especially useful to those who are physically disabled and who regularly rely on others. The paper's remaining sections are structured as follows. In Sect. 2, the existing system is presented. Section 3 discusses the proposed system for the Third Glove Eye Module. In Sect. 4, results and discussion are presented. In Sect. 5, conclusions and future work are discussed.

2 Existing System

A method for end-to-end real-time text localisation and recognition without prior knowledge of the words to be identified has been proposed. The challenge of character identification and segmentation is formulated as a real-time sequential selection from a set of extreme regions. In the first stage, the features were used to assess the likelihood of each ER being a character. For the second stage, which entailed using more computationally costly attributes to boost classification accuracy, only the ERs with the highest probability locally were picked. Character regions were labelled using an Optical Character Recognition (OCR) classifier trained on fake fonts and a very effective clustering technique that groups ERs into text lines. Depending on the scenario, the system creates two distinct sorts of noises. The main disadvantage of this approach was that blind persons could not distinguish between sounds, and even if they could, it would take some time to identify the front barrier based on the sound. An object detection sensor has been included in the new devices, which can identify an item but not the distance between the person and the object. Wahidur et al. [4] introduced a laser cane visible in three directions to bypass the limitation above. The angle is parallel to the ground and 45° above it, with a sharp decline. Because this laser only operates when an object or obstacle enters its range, and because the range is so short, it can only be utilised in internal systems and not in outdoor systems. Now, they are launching voice assistants that merely inform the user of the sensor's output, such as distance, direction, and object name. Prathipa et al. [3] developed the system for visually impaired and blind people. The proposed system can find obstacles in all directions using ultrasonic sensors and automatically calculates the distance to reach the destination using speech processing techniques. Harum et al. [2] designed a new model of book readers for blind people by using smart IoT technology. A picture can be taken using the camera, and the book reader

can process the image using OCR software. IoT systems use speech processing technology to convert online text to voice. Blind people can benefit from this model. People can easily gain new knowledge related to their interests. A navigation-based mobile robot was developed by Fong et al. [1]. Existing a low-cost ultrasonic distance sensor, this system gives the best results in avoiding obstacles in varied positions and under different environmental conditions. The Ultrasonic blind walking stick assists blind people in walking more comfortably and confidently. An Ultrasonic sensor (US) is used to sense the distance from any obstacle, and an RF remote is used to find the stick remotely. But nowadays, technology has improved a lot with the innovation of new technologies. It provides a smarter way to live and enjoy their lives in a sitting place, but physically challenged people still have to depend on others to do any work. We developed an intelligent device called the IoT-based Third Eye Glove to address this issue, allowing blind and deaf people to move and reach out without using a stick.

3 Proposed System for the Third Glove Eye Module

Our system uses Ultrasonic sensors linked with a health monitoring network to detect impediments and monitor blind people's health. These are combined with buzzers and vibration motors to alert users of nearby impediments. The goal is to make the user autonomous and shield them from deadly barriers. We propose a smart health monitoring system and smart gloves for visually impaired people. By using this gadget, users can avoid using a cane. This invention will let blind people navigate without a stick. They may wear it as a fabric band, and it works precisely with little training. Gloves will have the hardware. When a user wears gloves and walks, sensors detect impediments, buzzers ring, and a vibrating motor vibrates. A vibrator will inform the user when an impediment approaches. Smart shoe users do not need to rely on others when travelling somewhere. The clever blind guide system is affordable and user-friendly. It is being made to help blind and visually impaired people avoid barriers that could kill them and stay independent. To help visually impaired persons avoid obstacles, we suggest a solution that includes a smart health monitoring system that uses hardware components like Ultrasonic sensors, Buzzer, Vibrating Motor, Servo Motor, NodeMCU, Bluetooth, and Power Supply is shown in Fig. 1. The software package used in this model is Embedded C Programming Language, Arduino IDEtem, and a pair of smart gloves.

3.1 Ultrasonic Sensor

The transmitter is on the left, whilst the receiver is right. The transmitter circuit employs an ultrasonic sensor to determine the distance between the top of the bottle

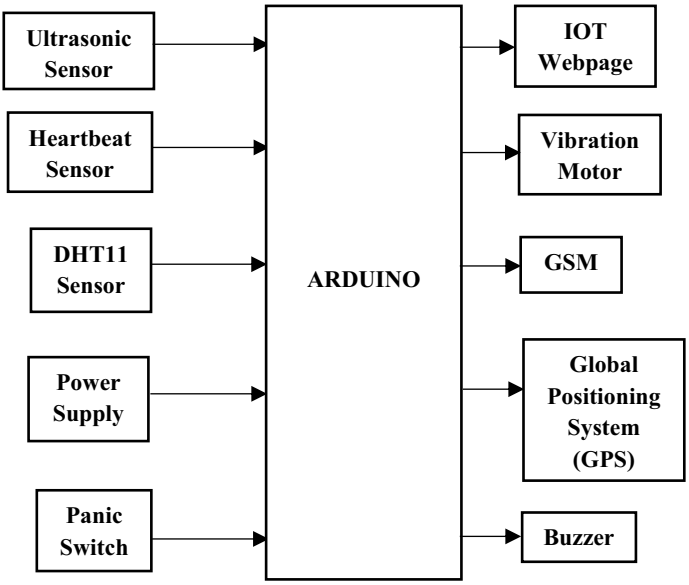


Fig. 1 Block diagram of the proposed system

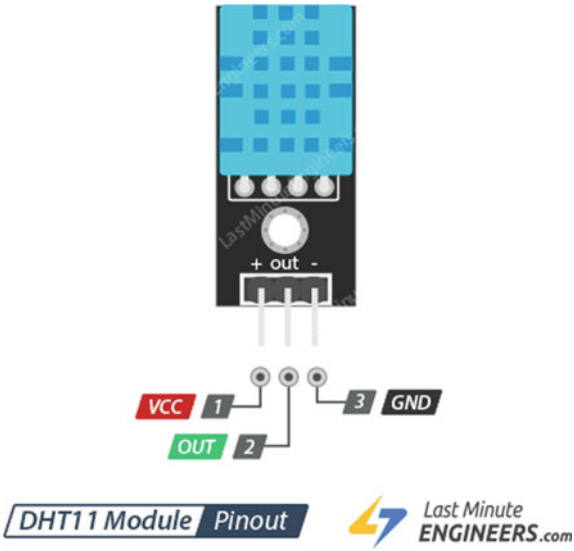
Fig. 2 Ultrasonic sensor

or tank and the water level. The Trigger and Echo apertures on an ultrasonic sensor are different. Figure 2 shows the Ultrasonic sensor.

3.2 DHT11 Sensor

The DHT11 is a low-cost digital sensor that measures both temperature and humidity is given in Fig. 3. It uses a capacitive humidity sensor and a thermistor to monitor the humidity in the air and then produces a digital signal to the data pin. Although it is simple to operate, accurate timing is required for data collection.

Fig. 3 DHT11 sensor



3.3 Heart Rate Sensor

Plug-and-play pulse rate sensing for Arduino and comparable boards, Pulse Sensor is a low-cost, space-saving option. Students, artists, athletes, manufacturers, and game and smartphone developers may all benefit from using real-time heart rate data in their work (Fig. 4).

Fig. 4 Heart rate sensor

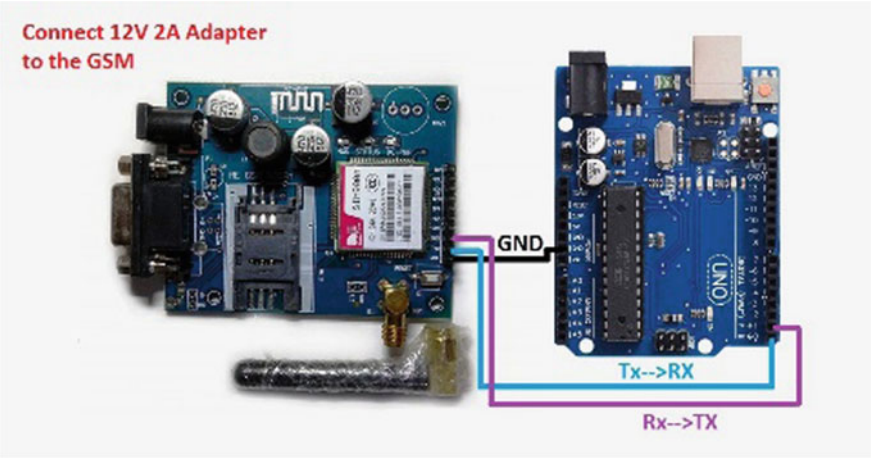


Fig. 5 Arduino-based GSM module

3.4 Gsm

The European Telecommunications Standards Institute produced the GSM (Global System for Mobile Communications) standard. It was designed to specify the protocols for second-generation digital cellular networks used by mobile phones and is currently the global standard for mobile communications, operating in over 220 countries and territories with a market share of over 90%. There are numerous GSM modules available on the market. For our project of connecting a GSM module to Arduino and sending and receiving SMS, it is usually advisable to select an Arduino-compatible GSM Module is shown in Fig. 5.

3.5 Gps

A GPS module may be integrated into IoT patient monitoring using Arduino. This GPS module will determine the patient's position or location based on the collected longitude and latitude information. The position will then be transmitted to the cloud, which is the IoT, via the Wi-Fi module (Fig. 6).

3.6 Voltage Regulator

The LM7812 is a TO-220 positive voltage regulator IC from the LM78xx family, produced by various electronic component manufacturers. The IC 7805 is a three-pin

Fig. 6 GPS chip

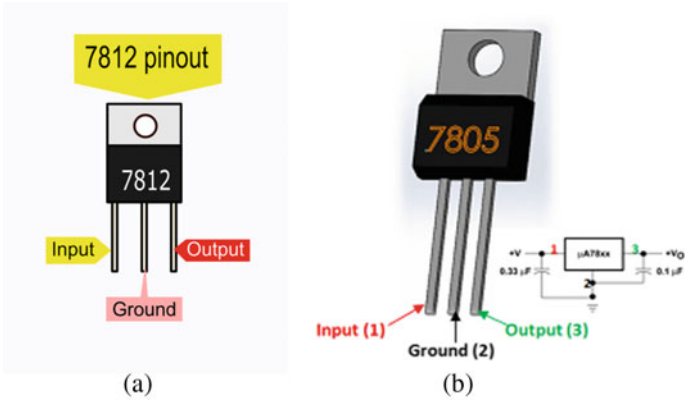


Fig. 7 **a** and **b** shows the pins of LM 7812, and IC7805

component with an input pin for DC voltage, a ground pin for the regulator's ground, and an output pin for 5 V. Fig. 7a and b shows the pins of LM7812 and IC7805.

3.7 IOT Output

The IoT refers to the billions of physical items linked to the internet and collects and shares data worldwide. It's now possible to turn everything, from a pill to a jet, into

an IoT component because of the development of low-cost computer processors and the widespread availability of wireless networks.

3.8 Smart Glove

In the user's gloves, the hardware will be corrected. When the user puts on the gloves and goes someplace, sensors on the hardware detect obstructions, and buzzers and vibrator motors sound. A vibrator will also warn the user when an impediment is approaching. When travelling, the wearer of the intelligent shoe does not need to rely on others. The technology is both affordable and user-friendly as a smart blind guide system. It is being developed to assist blind and visually impaired individuals in avoiding obstacles that can be fatal in many scenarios whilst still allowing them to be independent.

4 Results and Discussion

An ultrasonic sensor is used to detect the obstacle, and it alerts the person through the beep sound and vibrator motor in Fig. 8a. If an object is at a long-distance means, the sound will be beep beep. If the same for the object comes closer, then the sound will be like a beep. beep. beep. beep. beep., similarly to a vibrator motor. A panic switch is pressed to send the live location of the blind person to the caretaker via GSM location, as shown in Fig. 8a and b. The complete module of Third Eye Glove is depicted in the Fig. 9. Similarly, the DHT11 sensor senses temperature, humidity, and heartbeat sensor will give a heartbeat. Those values are updated on the IoT page in Fig. 10a.

Finally, the data recognised and fetched from all the sensors interfaced to the NodeMCU will be monitored on a separate page, i.e. an IoT page as shown in Fig. 10a, which helps the user constantly monitor their health, thus serving as an assistant in real life and also it shows some random values of a person during regular intervals of time includes temperature, humidity, and Heartbeat. It also includes an average of three parameters in the column shown in the Fig. 10a.

Figure 10b and 11a and b shows the graphical representation of temperature, Heartbeat, and humidity values taken at regular time intervals. And if there is any change in the Heartbeat value, an SMS will be sent to one of the family members.

5 Conclusions and Future Work

As part of the proposed project, we established two different systems. The first and second are smart gloves and a smart health monitoring system. We integrated the

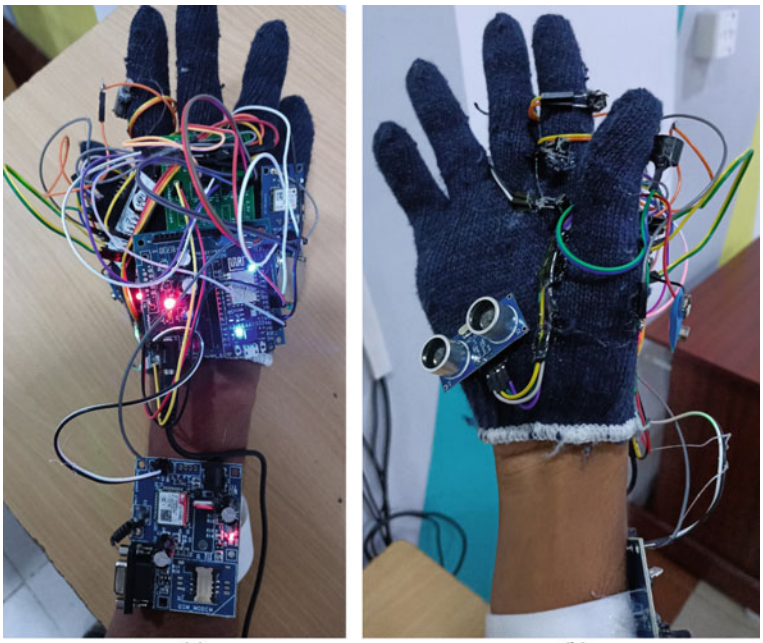


Fig. 8 **a** Front end of the third eye glove and **b** back end of the third eye glove

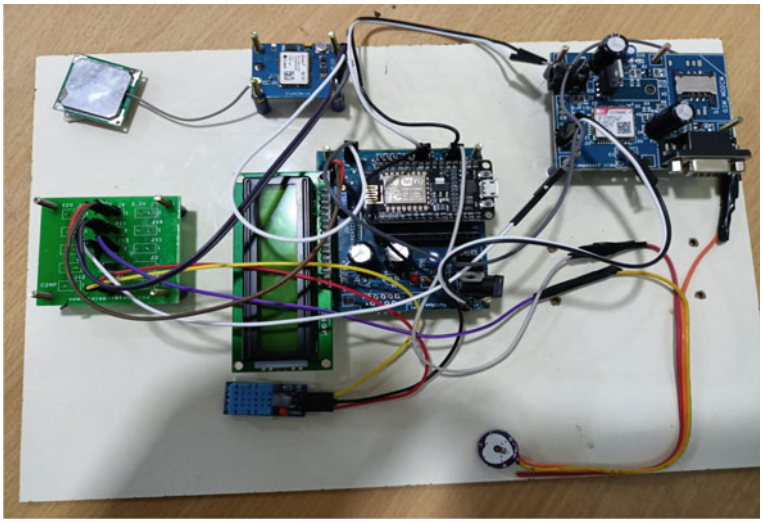


Fig. 9 Shows the third eye design module

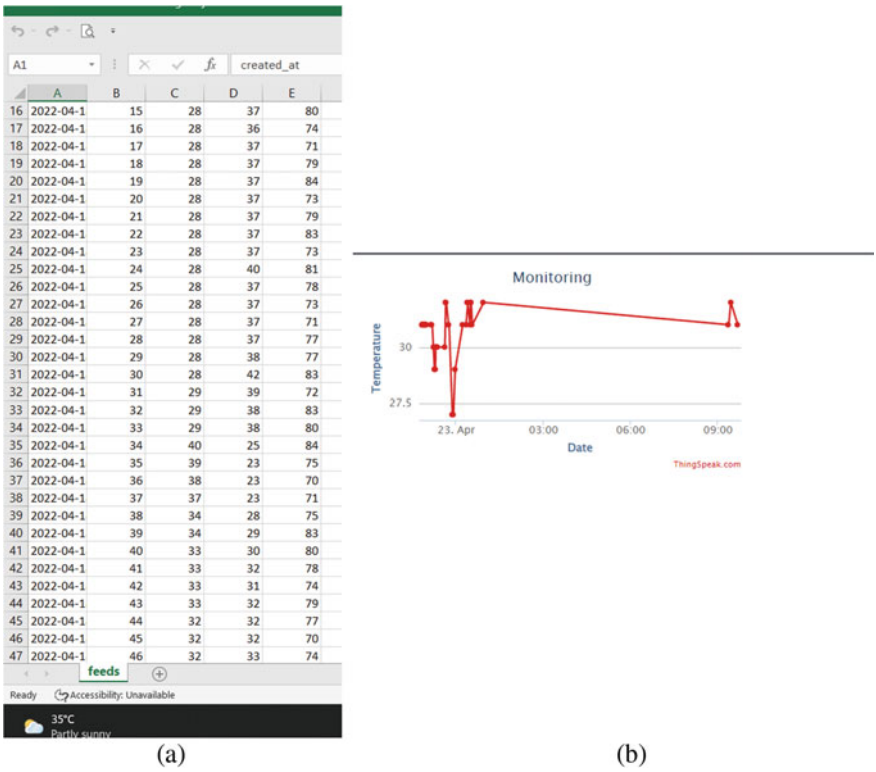


Fig. 10 **a** Parameters of third eye glove with IoT page, **b** temperature graph

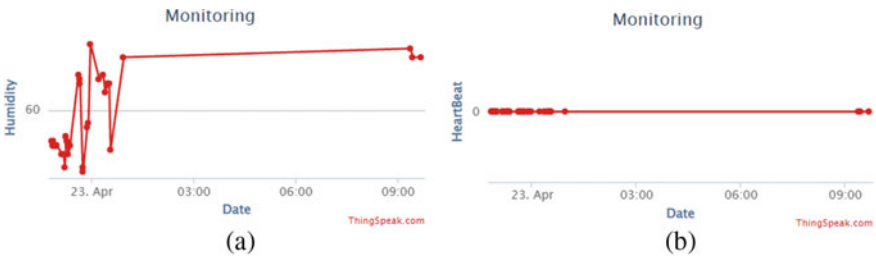


Fig. 11 **a** and **b** Shows the graphically plot for humidity, and heartbeat graph

two modules and made them work together to address past system flaws, including low coverage, energy problems, and false alarms, amongst others. This system used piezoelectric panels to generate energy whistle the user walks, eliminating the need to charge or replace batteries regularly, which is a prevalent issue in IoT technology. If one of the systems fails, the second system will continue to work independently and finish the task of finding obstacles without interruption. The proposed

system connects modules using wireless communication protocols such as Bluetooth and Zigbee, decreasing the number of cords. This device will immensely help the blind since it will allow them to commute indoors and outdoors without assistance, becoming self-sufficient and requiring no technical skills to operate. There will be many ways to expand the system in the future, and it will be easy to add more sensors or other parts.

References

1. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3–4):143–166. https://doi.org/10.1016/s0921-8890(02)00372-x
2. Harum NB, Azma N, Akmar N, Ayop Z, Anawar S (2019) Smart book reader for visual impairment person using IoT device. Int J Adv Comput Sci Appl 10(2). https://doi.org/10.14569/ijacsa.2019.0100233
3. Prathipa R, Arun M, Sownndarya C, Nivetha V (2018) Ultrasonic waist-belt for visually impaired person. Int J Emerg Technol Comput Sci Electron 26(4):23–25
4. Wahidur M, Islam R, Harun-Ar-Rashid M (2018) IoT based blind person's stick. Int J Comput Appl 182(16):19–21. https://doi.org/10.5120/ijca2018917824

Multimodal Peripheral Alert to Improve Teaching-Learning for Blended Classroom

Ujjwal Biswas and Samit Bhattacharya

Abstract Designing peripheral warnings and notifications to support teaching-learning is progressively increasing. However, existing work usually fails to effectively integrate real-time alerts and tackle poor students in a blended classroom. We present an in-class multimodal alert method for teachers and students to address the challenges. The system utilizes performance prediction and classification of students for real-time alert. The classification of students based on course performance helped in optimizing the number of alerts. The peripheral device selection aided in preventing the disruption in the lecture follow. Moreover, alert content delivery timing (start, during, and end of the class session) is used to reduce alert fatigue. We reported the design and the initial study results. The results show that 25 teachers and students reacted positively to the system design, technology, and features.

Keywords Peripheral alert · Real-time feedback · E-learning tool · Blended class

1 Introduction

Technology-driven teaching-learning has become popular with the rise of ICT [1, 2]. The bring-your-own-device (BYOD) paradigm enables students and teachers to support the technology-driven tools in personalized learning using ICT [2]. A blended learning platform combines the valuable features of online learning with face-to-face learning using both ICT and BYOD paradigms [3]. However, teachers cannot care for individual students in a classroom. Therefore, early warning and feedback systems have become a popular research field to take care of individual students [4, 5]. Such a technology-enhanced environment can aid the teachers and students in improving interaction and engagement in teaching-learning [6].

Understanding each student's performance state and providing appropriate alerts by the teachers is challenging during a lecture in a classroom. Existing studies show feedback reduces student failure and dropout, which in turn improves class-

U. Biswas (✉) · S. Bhattacharya
Indian Institute of Technology, Guwahati, Assam, India
e-mail: ibiswas.ujjwal@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_70

room learning outcomes [4]. Instructors are required to reflect on their performance achievement gaps, and performance improvement for a better learning outcome [7]. However, most of the teachers cannot do so due to their busy schedules. Teachers have multiple crucial duties, such as productive use of time, managing the classroom, and effective lecture delivery [4]. Teaching-learning at higher education levels is challenging to identify whether students are engaged and involved. The digital design of real-time alerts is vital to address the challenges. We have preferred to focus on such a system that has been ignored in state-of-the-art research [4].

The automated alert system is essential to manage classes effectively to enhance teaching-learning [9]. However, the effective integration of the students' academic performance metrics to identify at-risk students in a course and give both the students and teacher feedback is challenging. But the real-time feedback helps to learn from the mistakes of both instructor and students. There is a lack of real-time decision-making and motivation to enhance teaching-learning using timely warning and notification at the institutions level.

We reported an alerting system design that includes complex alert scheduling during lecture delivery, both for teachers and students. The system design is real-time feedback-driven for the blended classroom system. The system alert teachers when students are at risk, not engaged in learning, and not participating in real-time classroom discussions. The system will take care of alert fatigue and disturbances in lecture flow due to the alert. The alert helps the instructor evaluate himself and change their teaching pattern to make the class more engaging. Therefore, the solution aids the teacher and students with a digital alert to monitor the students' activity and provide feedback automatically.

2 Related Work

Educational research has focused on the design of e-learning tools to supporting teachers and students [10–12]. The use of digital devices such as smartphones, tablets, and personal digital assistants (PDAs) has become an integral part of our everyday teaching-learning [8]. However, during teaching, a teacher needs to continuously perceive each student's current performance state to support students [13]. Teachers feel escalating pressure to secure that all of their students succeed in the course. To address the challenges, smart classroom system has become popular.

Tools such as learning analytics (LA), students dash, and educational data mining (EDM) have been used to develop various kinds of smart classroom systems [14, 15]. These tools are popular for different purposes, such as characterizing students' performance [20], identifying their needs, and monitoring them [19]. A popular technology-enhanced classroom system handles a massive volume of data related to learners present in a classroom using multiple sensors, and devices [16, 17]. The sensory data assists in distinguishing whether a student engages with teaching-learning or not. Occasionally, the data is utilized to enhance students' level of understanding [16, 18]. To gain more meaningful use of those data demands effective decision-

making instruction in a day-to-day classroom activity. One of the solutions is to create new technology-enhanced e-learning tools to address the challenges of alerting and giving feedback to support teachers and students in a blended classroom [6].

In this direction, there is a lack of studies to give feedback to enhance teaching-learning by timely alert in a peripheral device. However, exceptional support in a classroom environment can potentially improve learning outcomes. Given these shortcomings, we introduced a alert system to explore how automated feedback can help teachers and students to address the above issues.

3 Proposed Alert System

The system allows sending alerts to the user across two devices using in-class Wi-Fi. The primary device is a smartphone or tablet. The secondary device is a peripheral smart watch/band. The system selects primary or secondary devices dynamically in real time to alert. The system uses three states (C: critical, LC: likely to be critical, and N: normal) similar to the work [19]. Our system will predict the state based on the students' course internal assessment data (e.g., quiz, attendance, and class test performance). Some students are performing well in all the aspects and require very less intervention, and we term these students to be having N type. There might be some students for whom the intervention is desirable, these are the LC types. There can also be students who must be given special attention, and we call them the C types. The states help to send an optimized number of alerts to students.

The basic idea and the logic flow of our proposed system are shown in Fig. 1. The system automatically determines actual alert timing for teachers and students using students' states and teachers' interaction data. The embedded sensory data (e.g., screen lock) of the smartphone/tablet is responsible for the teacher alert timing (see Fig. 1 right part). The teachers' device selection triggers a real-time alert. The teachers' device selection ensures that the teachers' lecture delivery should not hamper. Also, the system partitioned alert delivery based on timing (start, during, and end of the class session) to reduce human users' cognitive load for both students and teachers.

We assumed the number of students in the class is n, the machine learning algorithm is support vector machine (SVM), and three state classifier of the students (C, LC, and N). There are 16 alerts for both the students and teachers. To classify students, we need run-time complexity $O(s \times d)$, where dataset = $\{M, N\}$, dimension of M is n $\times$ d with d internal assessment data dimensionality, dimension of N is n $\times$ 3, and s are the number of support vector. After classification, we alert students by checking 16 alerts, so complexity to alert students is $O((16 + O(s \times d)) \times n)$. System alerts teachers by checking all the 16 alerts given to students, so it need $O(16 \times n)$. Therefore, the overall system complexity $O((16 + O(s \times d)) \times n) + 16 \times n)$.

In this article, we study the alert system to use in a real-time classroom. Here, we considered a class of an hour duration. Our system alerts C category students every

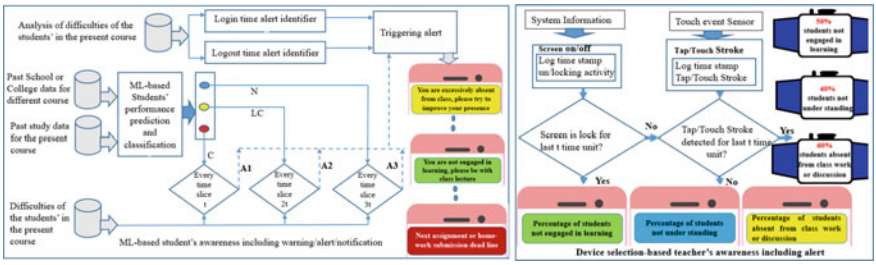


Fig. 1 Illustrates the design of the proposed system, the students' alert idea and the scheme (left) and the idea of peripheral alert to reduce lecture flow disturbance (right)

t minutes, LC students every $2t$ minutes, and N category students every $3t$ minutes, where $t = 5$ minutes (see in Fig. 1 left). In a one-hour lecture session, the worst-case alerts count is 12 for C, 6 for LC, and 4 for N state students. The alert count is not constant. The teacher can change the value of t according to their classroom handling experience. Theoretically, if we decrease the value of t, it will increase the alert count. In this case, a student may become alert fatigue. When we increase the value of t, it will decrease the number of alert counts. This will affect the positive side of the alert system.

4 Empirical Study and System Validation

We build a paper-based prototype for the empirical study. The goal of this study is to test alert scheme, system functions, and possible usability aspects. The following sections illustrate the prototype design and user perceptions about the system.

4.1 Prototype Design

The prototype helps to get feedback from the user about the system design. We reported the user satisfaction and user-friendliness of the innovative alert interface. User satisfaction testing is needed to ensure that real-time interventions are valuable and satisfactory to users. The design has the potential to enhance the access and acceptability of real-time alerts. We carry out these studies to ascertain the utility of the tool for teachers and students with real-time alert and feedback to improve further the interface design using controlled trial (see paper-based prototype as in Figs. 2 and 3).

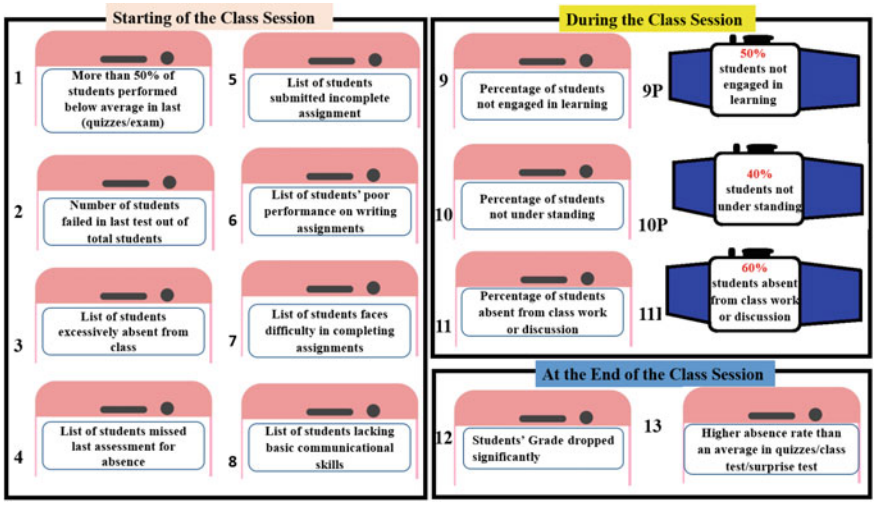


Fig. 2 Alerts used in the user study for teachers participants

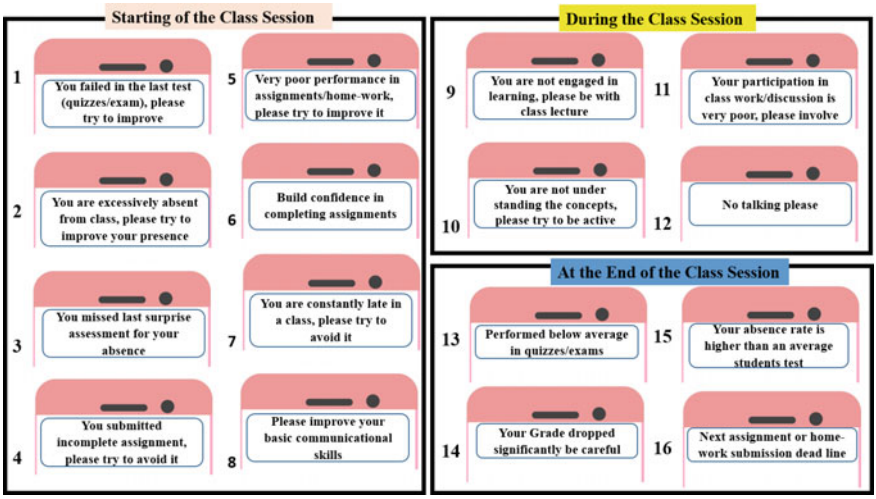


Fig. 3 Alerts used in the user study for student participants

4.2 User Study

We conducted an interview and feedback sessions with the user (teachers and student participants) in the Indian Institute of Technology, Guwahati.

Participants We collected data from a total of twelve "teacher" participants (eight male and four female) and thirteen "student" participants (nine male and four female).

The teachers' age group was 26–40 years, with an average age of 32.5 years, teaching experience of 6.5 years, and smartphone/tablet usage experience of 7.17 years. The students' age group was 26–36 years, with an average age of 29 years, interface design experience of 7.31 years, and smartphone/tablet usage of 7.77 years.

Procedure We explained the basic design idea of our system to the participants (both for teachers and students). We start with a short introduction, and signing the consent form takes about 10–15 minutes, including addressing their queries about the system functions. All the discussion audios are recorded to improve system design and post-analysis of the feedback. We requested users to provide their opinion on a printed copy of the questionnaires in a 5-point Likert scale about the system functions, including additional advice (see Table 1). We collect ratings for analysis and validation of functional requirements and their usefulness in real-time classroom use. Furthermore, we also include the seven yes/no construct items for teachers (TOQ:teachers open-ended question) and eight for student participants (SOQ:students open-ended question), including the opinion mentioned in the table (see Table 2). Each participant spent about 40–50 minutes of their valuable time in completing the entire session.

Results The above studies show that the acceptance of smartphone and wearable technologies in classroom alerting and feedback is exciting. Teachers rated the system functions with more than 75% (9 out of 12) cases with mean ratings between 4.17 and 4.67 and remaining between 3.50 and 3.67 (see Table 1 3rd column from the left). Also, for students participants, we received an issue where user mean ratings are below 4 (3.87) (see Table 1, 4th column from the left). Moreover, we found that both the users paid no attention to accepting and learning how to use smartphones and wearables (smart bands/watches) in the classroom. All teachers and students were satisfied with the novel idea and technologies to utilize it for real-time in-class teaching-learning. Although all the users participated in the studies, they are familiar with using smartphones and smart-band.

The qualitative tests analysis is based on 9 experienced teachers open-ended questionnaires (see Table 2). Most of the participants agreed on the system design features on TPQs statements. In TPQ6, 100% (9 out of 9) users reported that the intelligent peripheral alert features will reduce alert fatigue. In TPQ2, 83.33% (8 out of 9) users agreed that the list of alert and awareness features are helpful for in-class use. For other reaming feature sets (TPQ statements), 80.67% users agreed that the list of features is also crucial for real-time classroom use.

The qualitative tests analysis shows 13 students reported positive feedback in open-ended questionnaires (see Table 2). Most of the participants agreed on the system design features on SPQs statements. In SPQ1 to SPQ4 and SPQ6, 92.30% (12 out of 13) users decided that the intelligent features are useful in classroom teaching-learning. In SPQ5, 84.62% (11 out of 13) users agreed that the time gap t to $3t$ alert scheme is helpful for in-class use. For other reaming feature sets (SPQ7 and SPQ8 statements), 61.54% users agreed that the list of alert and awareness features is also essential.

Table 1 Statements for functionality and feature testing for users (teachers and students) perception about the proposed system interface

Teachers' perception statements about the system	Students' perceptions statements about the system	Teachers' rating		Students rating	
		Mean	SD	Mean	SD
Do you feel partitioning alert and awareness into three intervals (start of the class, during the lecture delivery, and end of the class session) reduce teachers' alert fatigue?	Do you feel partitioning alert and awareness into three intervals (start of the class, during the lecture delivery, and end of the class session) reduce students' alert fatigue?	4.67	0.471	4.23	0.890
Do you think the alert to teachers using a combination of smartphone and smart watch/band during the lecture delivery reduces teachers' alert fatigue instead of the warning using smartphone only?	Do you think the alert to students using a combination of alert type (text, audio, popup, and screen flash) during the lecture delivery reduces students' alert fatigue instead using notification only?	4.25	0.595	4.15	1.099
Do you think alerting teachers using a combination of smartphone and smart watch/band during the lecture delivery reduces users' technological distractions in lecture flow instead of using a smartphone only?	Do you think the alert to students using a combination of smartphone and smart watch/band during the lecture delivery reduces users' technological distractions in learning flow instead of using a smartphone only?	4	0.816	3.77	1.250
What is your opinion about students' characterization based on their performance metric and delaying alert for relatively good students will reduce alert fatigue?	What is your opinion about students' characterization based on their performance metric and delaying alert for relatively good students will reduce alert fatigue?	3.92	0.759	4.23	0.576
What is your opinion concerning the delaying alert t to $3t$ time interval that will reduce alert fatigue for relatively good students?	What is your opinion concerning the delaying alert t to $3t$ time interval that will reduce alert fatigue for relatively good students?	3.92	0.759	4.23	0.576
What is your view about the characterization of students based on their performance metric and delaying alert for relatively good students will optimize technological destruction?	What is your view about the characterization of students based on their performance metric and delaying alert for relatively good students will optimize technological destruction?	4.33	0.623	4.39	0.625
What is your view concerning the delaying alert t to $3t$ time interval that will reduce technological destruction for relatively good students?	What is your view concerning the delaying alert t to $3t$ time interval that will reduce technological destruction for relatively good students?	3.58	0.953	3.54	1.009
What is your perception regarding the delaying alert time interval t = fixed (may be 5 minutes) to reduce real-time alert fatigue?	What is your perception regarding the delaying alert time interval t = fixed (5 minutes) to reduce real-time alert fatigue?	3.85	0.593	4.24	0.678
The various functions for alert are well integrated for real-time large classroom use.	The various functions for alert are well integrated for real-time large classroom use.	4.33	0.623	4.23	0.697
The various functions for alert are necessary for real-time large classroom use.	The various functions for alert are necessary for real-time large classroom use.	4.53	0.824	4.62	0.667
I would like to use this alert in my classroom teaching.	I would like to receive this alert content in my classroom for further assistance and better learning.	4.67	0.471	4.08	0.917
I feel the system will help in customizing the lecture.	I feel the system will help to avoid future difficulties.	4.37	0.371	4.68	0.718

Table 2 Open-ended statements for users (teachers and students) qualitative analysis

Item No.	Teachers' open-ended questionnaire (TOQ)	Item No.	Students' open-ended questionnaire (SOQ)
TOQ1	Do you agree that the 16 alert mentioned above are useful for classroom teaching?	SOQ1	Do you agree that the 16 alert mentioned above are useful for classroom teaching-learning?
TOQ2	Do you agree that the 16 alert timing are useful for classroom teaching?	SOQ2	Do you agree that the 16 alert timing are useful for classroom teaching-learning?
TOQ3	Do you agree that the 16 alert timing are helpful to enhance your teaching?	SOQ3	Do you agree that the 16 alert and their timing are helpful to enhance your learning?
TOQ4	Do you agree that the three real-time (during class) peripheral alert is useful for classroom teaching?	SOQ4	Do you agree that the four intervals t to $3t$ time interval (critical category students to normal students) vary for real-time warning are useful for classroom teaching-learning?
TOQ5	Do you agree that the three real-time (during class) peripheral warning will reduce class flow disturbance?	SOQ5	Do you agree that the four intervals t to $3t$ time interval (critical category students to normal students) vary for real-time alert will reduce class flow disturbance?
TOQ6	Do you agree that the three real-time (during class) peripheral alert will reduce warning fatigue?	SOQ6	Do you agree that the four intervals t to $3t$ time interval (critical category students to normal students) vary for real-time warning will reduce warning fatigue?
TOQ7	Do you agree that the alert and awareness messages are sufficient for real-time classroom use?	SOQ7	Do you agree that the time interval t = 5 minutes is sufficient for real-time alert to reduce fatigue?
		SOQ8	Do you agree that the alert and awareness messages are sufficient for real-time classroom use?
Additionally, if you want to add some functionalities or features in smart alert design, please mention those			

5 Discussion and Design Implications

This manuscript has reported peripheral alert methods to facilitate in-class teaching-learning and empirical validation of the system design. We suggested the three predicted states. However, that might be challenging for universities, particularly those with small student records. We utilize the classification of three states [19] in our alerting and peripheral selection to avoid alert fatigue in real-time face-to-face teaching-learning, which is novel and unique. Teachers and students confirm that

alert is crucial for better learning outcomes. The idea is novel and valuable based on user feedback (see Table 1).

The results show that teacher and students participants preferred the system design (see Table 1). The user found the tool helpful for real-time classroom teaching-learning. All teacher participants agreed that the alert and peripheral feedback would reduce alert fatigue. Moreover, 10 out of 12 participants believe that one crucial design goal, lecture flow disturbance, can be optimized using teachers' peripheral selection. The teachers prefer characterizing and classifying students into performance states to optimize the number of alerts during the lecture. On the other hand, student participants also appreciated and gave positive comments on the tool's uses. Users also agreed that only smartphone alert content is non-persistent, as there are more chances of disturbing the teachers' lecture flow. The persistence of alerting based on the teacher's peripheral alert adds to improve usability and helps to ensure alert fatigue established through the empirical study.

We obtained positive comments from participants in three directions usefulness, reduced alert fatigue, and disturbance in lecture flow (see Table 1). Experience users advised that combining more explicit alerts and feedback contents would be more helpful. The limitation of our system is determining the alert contents. It is challenging to compute some of the alert content, such as identifying students talking in a class and understanding the concepts from their mobile sensor. Some other contents are more straightforward to compute, such as the attendance, the quiz, and the class test performance. Our alert system is not focused on these challenges. However, this is not a limitation because our system still uses whatever alert contents are available.

6 Conclusions and Future Work

The multiple conditions, features, and design goals make the alert system innovative and unique. Developing peripheral alerts for real-time classroom users is challenging due to difficulties getting users for testing in the initial stage. The prototype-based studies were performed to overcome the challenges. The study confirmed that the innovative alert methods are suitable to engage the teacher and student for better learning outcomes. Teachers and students agreed that systematic alerts and feedback would motivate them to rectify themselves when they regularly know their performance.

We hope that the findings and system design considerations presented will encourage future researchers in the educational community. We hope handling some of the research challenges and hurdles inherent in designing the system should motivate HCI and academic researchers in the future. We are currently gathering additional user data to refine and validate the suggested alert methods and techniques. We focus on identifying and including all multimodal alerts that influence the real-time alerts. We are also developing a framework to comprehend alerts better and be usable in a blended classroom.

References

1. Choudhury N, Venkatesh T, Bhattacharya S, Sarma S (2016) Avabodhaka: a system to analyse and facilitate interactive learning in an JCT based system for large classroom. Proc Comput Sci 84:160–8
2. Choudhury N, Tamarapalli V, Bhattacharya S (2015) An ICT-based system to improve the learning experience in a large classroom. In: 2015 IEEE seventh international conference on technology for education (T4E). IEEE, pp 27–30
3. Tikadar S, Bhattacharya S, Tamarapalli V (2018) A blended learning platform to improve teaching-learning experience. In: 2018 IEEE 18th international conference on advanced learning technologies (ICALT). IEEE, pp 87–89
4. Martinez-Maldonado R, Clayphan A, Yacef K, Kay J (2014) MTFeedback: providing notifications to enhance teacher awareness of small group work in the classroom. IEEE Trans Learn Technol 8(2):187–200
5. Wang Z, Zhu C, Ying Z, Zhang Y, Wang B, Jin X, Yang H (2018) Design and implementation of early warning system based on educational big data. In: 2018 5th international conference on systems and informatics (ICSAI). IEEE, pp 549–553
6. Rodríguez-Triana MJ, Prieto LP, Vozniuk A, Boroujeni MS, Schwendimann BA, Holzer A, Gillet D (2017) Monitoring, awareness and reflection in blended technology enhanced learning: a systematic review. Int J Technol Enhanced Learn 9(2–3):126–150
7. An P, Bakker S, Ordanovski S, Taconis R, Paffen CL, Eggen B (2019) Unobtrusively enhancing reflection-in-action of teachers through spatially distributed ambient information. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–14
8. Abugohar M, Yunus K, Rabab'ah G, Ahmed T (2019) Integrating cloud world synergy in ELT to adults: perspectives of handheld technologies
9. Singley MK, Lam RB (2005) The classroom sentinel: supporting data-driven decision-making in the classroom. In: Proceedings of the 14th international conference on World Wide Web, pp 315–321
10. Fortin PE (2019) Methods and interfaces for closed loop smartphone notifications. In: Proceedings of the 21st international conference on human-computer interaction with mobile devices and services, pp 1–5
11. Kushlev K, Proulx J, Dunn EW (2016) Silence your phones. Smartphone notifications increase inattention and hyperactivity symptoms. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 1011–1020
12. Ketabdar H, Moghadam P, Roshandel M (2012) Pingu: a new miniature wearable device for ubiquitous computing environments. In: 2012 sixth international conference on complex, intelligent, and software intensive systems. IEEE, pp 502–506
13. Braun MC, Beuck S, Wölfel M, Scheurer A (2018) Investigating multimodal warnings for distracted smartphone users on the move in potentially dangerous situations. Transactions on computational science XXXII. Springer, Berlin, Heidelberg, pp 1–14
14. Holstein K, McLaren BM, Aleven V (2017) Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms. In: Proceedings of the seventh international learning analytics and knowledge conference, pp 257–266
15. Aldowah H, Al-Samarraie H, Fauzy WM (2019) Educational data mining and learning analytics for 21st century higher education: a review and synthesis. Telemat Inf 37:13–49
16. Talwar J, Ranjani S, Aras A, Bedekar M (2013) Intelligent classroom system for qualitative analysis of students' conceptual understanding. In: 2013 6th international conference on emerging trends in engineering and technology. IEEE, pp 25–29
17. Augusto JC, Nakashima H, Aghajan H (2010) Ambient intelligence and smart environments: a state of the art. Handbook of ambient intelligence and smart environments, pp 3–31
18. Mathioudakis G, Leonidis A, Korozi M, Margetis G, Ntoa S, Antona M, Stephanidis C (2014) Real-time teacher assistance in technologically-augmented smart classrooms. Int J Adv Life Sci 6(1):62–73

19. Bhattacharya S, Shah VB, Kumar K, Biswas U (2021) A real-time interactive visualizer for large classroom. ACM Trans Interact Intell Syst (TiiS) 11(1):1–26
20. Tikadar S, Bhattacharya S (2021) Detection of affective states of the students in a blended learning environment comprising of smartphones. Int J Human-Comput Int 37(10):963–980

Sentinel-2 Data Processing for Pichavaram Mangrove Forest Using Convolutional Neural Network

S. Sudharsan, R. Surender, and Nandini G. Iyer

Abstract Image classification is commonly utilized in computer vision tasks like remote sensing, scene analysis, object detection, and image retrieval. We use the Sentinel-2 satellite dataset in our study to classify land cover using the convolutional neural network (CNN) method and determine different plant indices, water indices, and geology features in the Pichavaram mangrove forest. Mangrove woods protect the shoreline from ocean waves, Tsunami storms, and soil erosion. They are efficient at sequestering and storing carbon and mitigating climate change. It is critical to map the extent of the mangroves to protect them. We need an automation solution because the geographic expansions are so large. For automatic feature extraction, CNN is used. Our datasets have been divided into water bodies, vegetation, and marshy plains. Sentinel-2 has a variety of uses, including land monitoring, yield prediction, land cover, flood volcanic eruption detection, and landslide detection.

Keywords Image processing · Convolutional neural network (CNN) · Sentinal-2 data

1 Introduction

Mangrove trees are essential because they provide a coastline that protects against ocean waves and Tsunami storms. It is important for mapping and figuring out where things are, finding water resources, farming, figuring out where wetland areas are, getting geological information, and making plans for urban and regional planning

S. Sudharsan (✉) · R. Surender · N. G. Iyer
Department of ECE, Rajalakshmi College of Engineering, Rajalakshmi Nagar, Thandalam, Chennai, Tamil Nadu 602105, India
e-mail: sudharsan.s@rajalakshmi.edu.in

R. Surender
e-mail: surendar.r@rajalakshmi.edu.in

N. G. Iyer
e-mail: nandhini.praveen@rajalakshmi.edu.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_71

management. However, the classification of remote sensing photos remains a time-consuming operation due to its complexity [1]. The use of remote sensing data is critical for identifying the area of mangrove forests and comprehending their dynamics. The operational satellites of Sentinel-2 have created a new methodology for remote sensing that combines high spatial and temporal resolution with several spectral bands [2]. As a result, we use a deep learning-based strategy for autonomous feature learning. CNN is a deep learning model with convolutional, pooling, and full connection layers. Because of its efficacy in spatial feature exploration, the CNN algorithm is useful for high-resolution picture categorization. So, using the CNN algorithm, the Pichavaram mangrove forest is mapped and categorized for future investigation. In addition, multispectral satellite photos are used to calculate the normalized satellite indices. These photos highlight a certain phenomenon while minimizing other elements that diminish the image's effects. Healthy vegetation, for example, in the index image, will be bright, while unhealthy vegetation will have lower values, and empty land will be dark. Because the shading of hills and valleys affects the intensity of photographs; the indices are made so that the color of an object is more important than its intensity or brightness.

The remaining sections of the paper are organized in the following manner. Section 2 presents a literature survey on CNN for image feature extraction. In Sect. 3, the proposed system for feature extraction using Sentinel-2 data is discussed. In Sect. 4, results and discussion are presented. In Sect. 5, conclusions are discussed.

2 Literature Survey on CNN for Image Feature Extraction

Rajesh et al. [1] suggested a method for classifying the cities of Madurai and Ahmedabad using the CNN algorithm to extract and categorize the features. Compared to randomly initialized filter values, the proposed system extracts deep features using pre-defined filter values, enhancing the overall performance of the process. In detailed satellite images, the object-based categorization method can maintain edge information. A study was also conducted to locate surface coal mines from satellite pictures using deep learning algorithms [3]. The VGG network was utilized in conjunction with transfer. The validation dataset of the pre-trained VGG architecture had a classification accuracy of 94%. Rai et al. [4] suggested a CNN technique for Landsat 8 OLI satellite image classification. Principal component analysis (PCA) is performed on fused images to reduce dimension. Following that, CNN was used to classify the data.

The suggested method's classification results are compared to FCM, GIFP-FCM, and FKLICM. Kadhim and Abed [5] used CNN for feature extraction with pre-training models AlexNet, VGG19, GoogLeNet, and Resnet50. On three independent datasets, SAT4, SAT6, and UC Merced Land. Resnet50 model gets a more promising result than other models by examining the geographical distribution of mangrove forests. The methods given by Iovan et al. [2] are based on a deep CNN and have been evaluated using Worldview 2 and Sentinel-2 images. The results are excellent, and

they open up the possibility of autonomously surveying the distribution of mangroves across vast distances.

3 Proposed System

In this proposed system, the image obtained from the Sentinel-2 satellite undergoes the first step of dataset preprocessing. The processed image is fed into the CNN layers and trained and tested in two phases. The output of the images is classified into six classes and analyzed with their natural parameters. Hence, the accuracy is calculated for the study of the Pichavaram mangrove forest. Figure 1 shows the proposed block diagram of the system.

Pichavaram mangrove forest information was gathered from the United States Geological Survey (USGS) Web site. Sentinel-2 records were obtained, and they contain 12 bands: coastal aerosol, blue, green, red, vegetation red edge, near infrared (NIR), narrow NIR, water vapor, short-wave infrared (SWIR), and SWIR-cirrus [6]. PCA is used to preprocess the images, producing 3D patches and partitioning the datasets into training and testing stages in a 70:30 ratio. The CNN method with numerous layers is then used to classify it. We utilized the ADAM optimizer for training, which iteratively updates the network weights. Early stopping is used to prevent the model from becoming overtrained. The accuracy and loss for the epochs are investigated using the results collected.

3.1 Data Augmentation

Data from Sentinel-2 may be acquired by visiting the Web site maintained by the USGS. The proposed work area chosen from the geographical location is Pichavaram mangrove forest. The images from the Sentinel-2 satellite are then chosen for the same area and downloaded in jpeg format.

3.2 Data Preprocessing

These datasets are preprocessed with SNAP, a program that does radiometric calibration and atmospheric correction to improve the data. We enhanced a 12-band spectrum dataset from the Sentinel-2 satellite image, as shown in Fig. 2. As a result, we employ red, green, and blue bands to visualize our data.

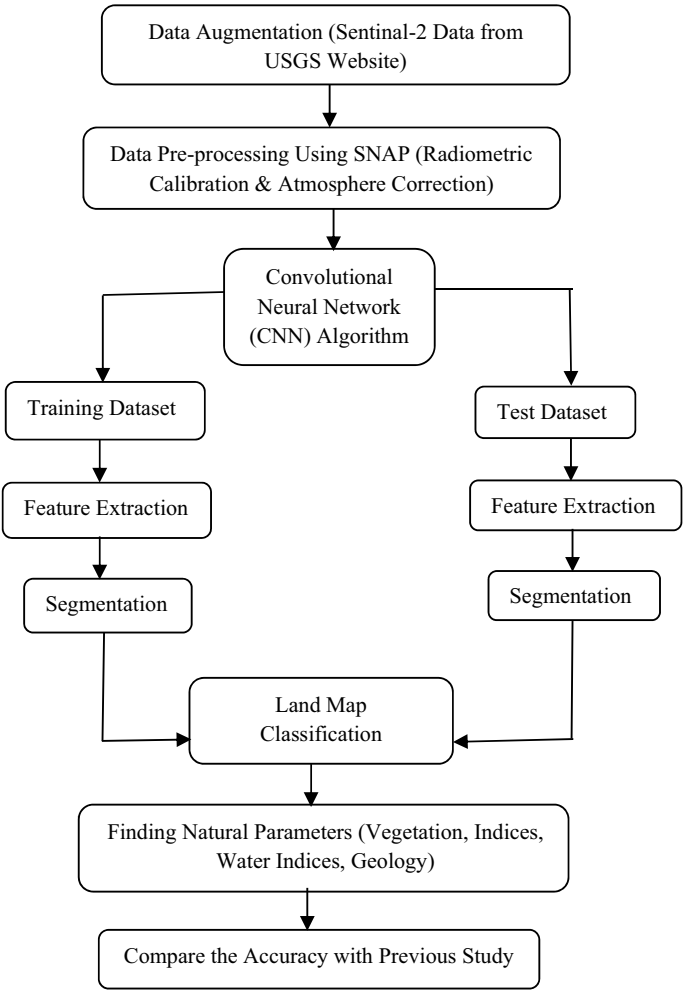


Fig. 1 Block diagram of the proposed system

3.3 CNN Algorithm

Because the 3D CNN requires three-dimensional data as an input, we must divide the satellite image into patches, each with its class. As a result, the photos are converted to ". mat" format. The patch's center pixel class has been defined as the patch's class label. Every patch has a specific size (W, W, and B). W and B are the window size and number of bands, respectively. Following that, PCA is implemented; 3D patches are created, and data are separated into train and test data in a 70:30 ratio. The CNN algorithm uses the input layer, followed by convolution layers with filter value, kernel size, and ReLU activation. The output layer is obtained when completely

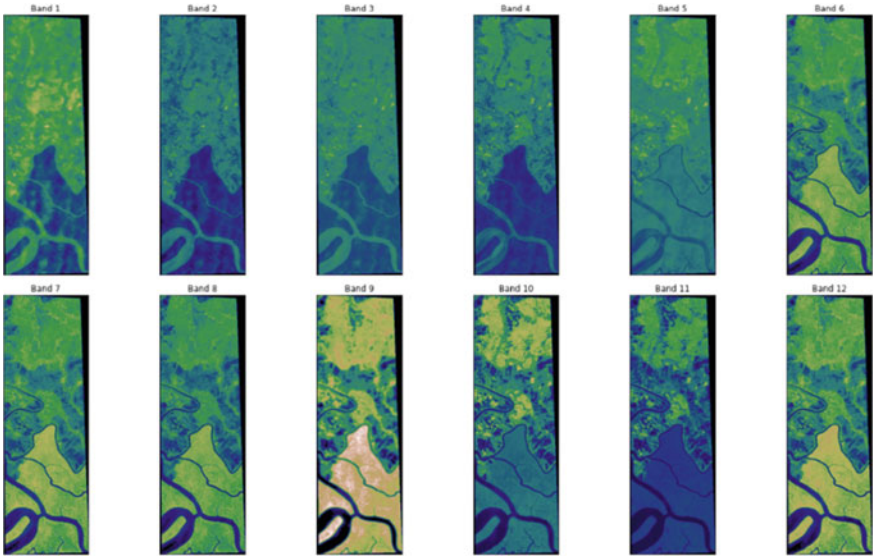


Fig. 2 Data preprocessing

linked layers with dense and dropout are established. The image's feature is extracted next, followed by segmentation.

3.4 Land Map Classification

As a result, the land map classification is divided into six categories: river, open scrubs, trees, lands, urban, and mangrove swamps.

3.5 Finding Natural Parameters

Biological parameters such as the normalized difference vegetation index (NDMI), soil adjusted vegetation index (SAVI), modified normalized difference water index (MNDWI), normalized difference moisture index (NDMI), clay mineral ratio (CMR), and ferrous mineral ratio (FMR) are discovered after the land map classification of Pichavaram mangrove forest.

3.6 Compare Accuracy

The training's accuracy and loss graph are then plotted using the matplot toolkit. The X-axis shows the epochs, while the Y-axis represents the percentage. The graph representation comes after the confusion matrix (CM). It demonstrates how rivers, open scrubs, forests, farmland, urban areas, and mangrove swamps are classified. We also displayed the Pichavaram mangrove forest classification report. The classification report includes accuracy, precision, recall, $F1$-score, and support for each class. The created CNN model's accuracy and CM are calculated, and the classification report is displayed.

4 Results and Discussion

4.1 Accuracy and Loss Graph of Training and Testing

Train accuracy improves when the number of epoch parameters is increased although train loss is larger in the first epoch and decreases as the epoch is adjusted. Test accuracy improves as the number of epoch parameters increases, but test loss is larger in the beginning and decreases as the period progresses. The graph depicting the precision and loss of training and testing datasets is shown in Fig. 3.

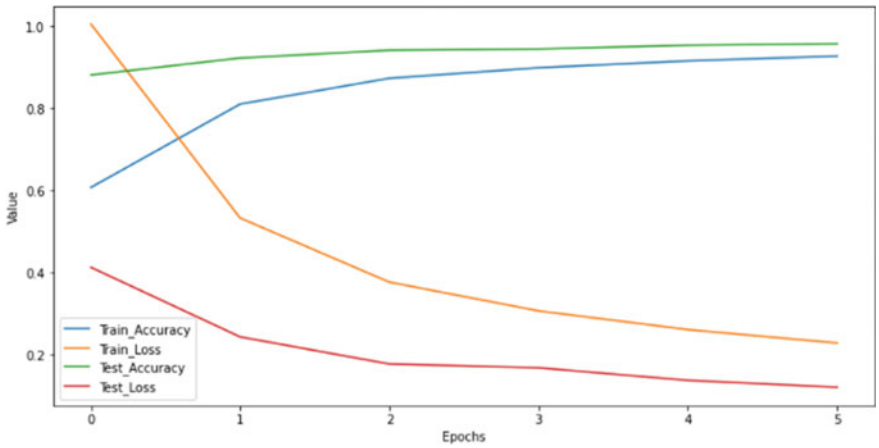


Fig. 3 Training and testing datasets' precision and loss graphs

Fig. 4 Confusion matrix prediction results

4.2 CM

A CM is a compilation of prediction results from six different classifications: river, open scrubs, trees, lands, urban, and mangrove swamps. From Fig. 4, an analysis of various classes has been computed. Class 1, representing the river, has an accuracy rate of 14,869. Class 2, representing the available scrubs, has an accuracy rate of 23,447. Class 3, representing the trees, has an accuracy rate of 6092. Class 4, representing the lands, has an accuracy rate of 16,138. Class 5, representing the urban, has an accuracy rate of 13,220. Class 6, representing the mangrove swamps, has an accuracy rate of 7765.

Table 1 shows the precision, recall, $F1$-score, and support for all classes like a river, open scrubs, trees, lands, urban, and mangrove swamps have been calculated with their values, respectively. $F1$-score has an overall accuracy of 0.96 and support of 85,288. Similarly, the accuracy, $F1$-score, support, and recall values of macro averages (MA) and weighted averages (WA) are computed and given in Table 1.

4.3 Natural Parameters

NDVI

The density of vegetation on a piece of land may be estimated using the NDVI, a metric that considers the amount of visible and NIR sunlight reflected by the plant. Figure 5a depicts the vegetation in three segments: dark green, light green, and yellow, each with a value of 1, 0.75, and 0.5. The calculated NDVI value was 0.756, corresponding to a lush green forest.

Table 1 Classification report on various classes

Class	Precision	Recall	*F*1-score	Support
River	0.96	0.92	0.94	16,222
Open scrubs	0.98	0.99	0.98	23,570
Trees	1.00	1.00	1.00	6095
Lands	0.96	0.96	0.96	16,790
Urban	0.96	0.98	0.97	13,545
Mangrove swamps	0.85	0.86	0.85	9066
The average accuracy of various classes				
Accuracy	–	–	0.96	85,288
MA	0.95	0.95	0.95	85,288
WA	0.96	0.96	0.96	85,288

$$\text{NDVI} = ((\text{NIR} - \text{Red})/(\text{NIR} + \text{Red})) \tag{1}$$

SAVI

SAVI is a vegetation index that attempts to reduce soil brightness effects by employing a soil-brightness adjustment factor. The dry areas with minimal flora are frequently found in SAVI. Figure 5b depicts the vegetation in three segments: dark green, light

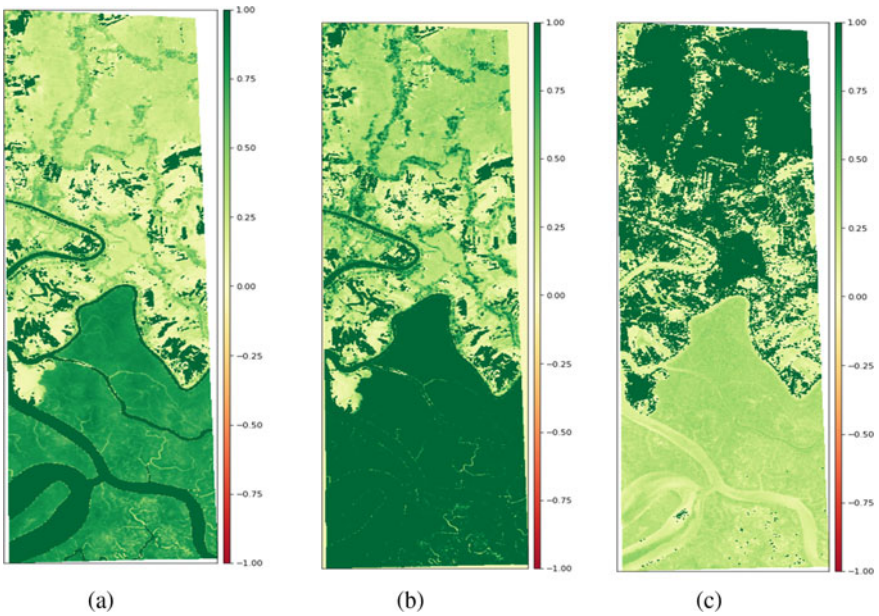


Fig. 5 **a–c** Natural parameters of NDVI, SAVI, and VARI

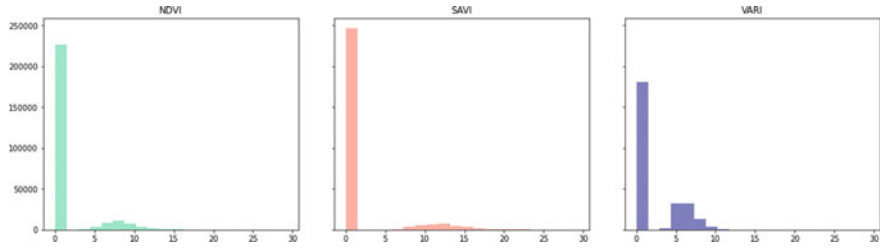


Fig. 6 Comparison of NDVI, SAVI, VARI

green, and yellow, each with a value of 1, 0.75, and 0.5. The calculated SAVI score was 0.865, which represents a lush green forest.

$$\text{SAVI} = ((\text{NIR} - \text{Red})/(\text{NIR} + \text{Red} + \text{L})) \times (1 + \text{L}) \tag{2}$$

VARI

VARI is designed to showcase vegetation in the visible spectrum while reducing lighting and atmospheric effects disparities. Because it utilizes all three color bands, it works well with RGB and color photographs. The VARI represents the blue and green colors in Fig. 5c. The higher light is scattered by the blue atmosphere, whereas the lower light is scattered by the green atmosphere. The computed VARI value was 0.125.

$$\text{VARI} = (\text{Green} - \text{Red})/(\text{Green} + \text{Red} - \text{Blue}) \tag{3}$$

The image's NDVI, SAVI, and VARI pixel distribution is depicted in Fig. 6. The high peak of the NDVI is approximately 23000 px; the high peak of SAVI is approximately 24000px, and the high peak of VARI is approximately 22000 px.

MNDWI

To enhance open water features, MNDWI employs green and SWIR bands. It also reduces built-up area characteristics, frequently linked to open water in other indices. The amount of water in the photograph is represented in Fig. 7a. It was discovered to be around 0.256.

$$\text{MNDWI} = (\text{Green} - \text{SWIR})/(\text{Green} + \text{SWIR}) \tag{4}$$

NDMI

NDMI is a way to measure how sensitive plants are to changes in moisture. It is used to keep track of droughts and fuel levels in fire-prone areas. It creates a ratio that mitigates illumination and atmospheric impacts by combining NIR and SWIR bands. Figure 7b represents the moisture content of the image. It was found to be around 0.334.

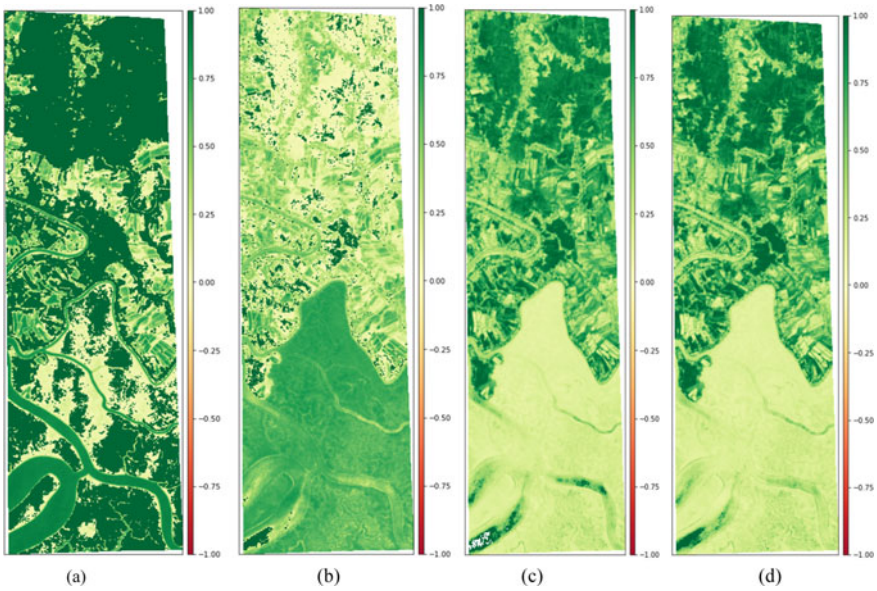


Fig. 7 **a–d** MNDWI, NDMI, CMR, and FMR

$$\text{NDMI} = (\text{NIR} - \text{SWIR1})/(\text{NIR} + \text{SWIR1}) \tag{5}$$

CMR

CMR is computed by multiplying together the SWIR1 and SWIR2 bands. This ratio exploits the fact that hydrous minerals such as clays and alunite absorb 2.0–2.3 microns radiation. Because it is a ratio, this index mitigates lighting changes caused by terrain. Figure 7c represents the clay mineral content of the image. It is found to be around 0.07.

$$\text{CMR} = \text{SWIR1}/\text{SWIR2} \tag{6}$$

FMR

The FMR highlights minerals containing iron-bearing elements. It takes advantage of the ratio between the SWIR and NIR bands. Figure 7d represents the ferrous mineral content of the image. It is found to be around 0.1.

$$\text{FMR} = \text{SWIR}/\text{NIR} \tag{7}$$

5 Conclusions and Future Work

The examination of the Pichavaram mangrove forest was successful. Using the CNN algorithm, the forest was split into six categories: river, open scrubs, trees, farms, urban, and mangrove swamps. There were a total of 618,796 images in the USGS collection. The ratio of instruction to testing was 70–30% (Training and Testing). Using the matplot package, the categorization map was created. After training, the CNN model achieves an accuracy of 95.60%. Traditional techniques support vector machine (SVM) and K-nearest neighbor classifier (K-NNC) have an accuracy of 93.80 and 92.46, respectively. The precisions for the river, open scrubs, trees, lands, urban, and mangrove wetlands are 0.96, 0.99, 1.00, 0.95, 0.96, and 0.86, respectively. The natural characteristics of the Pichavaram mangrove forest show that vegetation indicators, including the NDVI, SAVI, and VARI, are provided. The MNDWI and NDMI are then displayed to aid with water index visualization. In addition, the clay minerals ratio and the ferrous minerals ratio are computed and shown. These parameters are used to examine the forest.

References

1. Rajesh S, Nisia TG, Arivazhagan S, Abisekaraj R (2019) Land cover/land use mapping of LISS IV imagery using object-based convolutional neural network with deep features. J Indian Soc Remote Sens 48(1):145–154. https://doi.org/10.1007/s12524-019-01064-9
2. Iovan C, Kulbicki M, Mermet E (2020) Deep convolutional neural network for mangrove mapping. In: IGARSS 2020-2020 IEEE international geoscience and remote sensing symposium. https://doi.org/10.1109/igarss39084.2020.9323802
3. Shabbir A, Ali N, Ahmed J, Zafar B, Rasheed A, Sajid M, Ahmed A, Dar SH (2021) Satellite and scene image classification based on transfer learning and fine tuning of ResNet50. Math Probl Eng 2021:1–18. https://doi.org/10.1155/2021/5843816
4. Rai AK, Mandal N, Singh A, Singh KK (2020) Landsat 8 OLI satellite image classification using convolutional neural network. Procedia Comput Sci 167:987–993. https://doi.org/10.1016/j.procs.2020.03.398
5. Kadhim MA, Abed MH (2019) Convolutional neural network for satellite image classification. Intell Inf Database Sys Recent Dev 165–178. https://doi.org/10.1007/978-3-030-14132-5_13
6. Huang M, Jin S (2020) Rapid flood mapping and evaluation with a supervised classifier and change detection in Shouguang using Sentinel-1 SAR and Sentinel-2 optical data. Remote Sens 12(13):2073. https://doi.org/10.3390/rs12132073
7. Madhuanand L, Sadavarte P, Visschedijk A, Denier Van Der Gon H, Aben I, Osei F (2021) Deep convolutional neural networks for surface coal mines determination from sentinel-2 images. Eur J Remote Sens 54(1):296–309. https://doi.org/10.1080/22797254.2021.1920341

Machine Learning Prediction if the Patient is at Risk of Undergoing Surgery Based on Preoperative Medical Reports

Varsha Naik, Shakti Kinger, Ishaan Shanbhag, and Mufaddal Ragib

Abstract The ultimate aim of preoperative medical assessment is to minimize the patient's surgical and anesthetic preoperative morbidity or death, and to return him to normal functioning as soon as feasible. Any preoperative evaluation must begin with a thorough history and physical examination, with an emphasis on risk factors for cardiac and pulmonary problems as well as determining the patient's functional ability. We analyzed a real-world dataset of patients' blood tests by using machine learning techniques and data science. We had raw data in CSV format; we had to normalize the data and then analyzed the data and then visualized the data. After that we used different machine learning models on the data. We have used different libraries such as sklearn, pandas, scipy, numpy, and matplotlib. Among a large variety of algorithms, we have used by the included study; the highest accuracy we have obtained is 83.34% from logistic regression technique using random state. We have obtained close to similar result using decision tree with 81.81% accuracy. As far as our other algorithms used, we have obtained 73% accuracy and 80% by using gradient boosting classifier and support vector machine algorithms, respectively. The remaining algorithms that were not fruitful enough giving us −13.3% and −9% for gradient boosting regressor and linear regression, respectively.

Keywords Preoperative assessment · Perioperative risk · Preoperative preparation · Anesthetic risk · Logistic regression · Decision tree · Support vector machine

V. Naik (✉) · S. Kinger · I. Shanbhag · M. Ragib
Dr. Vishwanath Karad MIT World Peace University, Survey No: 124, Paud Rd, Kothrud, Pune, Maharashtra 411038, India
e-mail: varsha.powar@mitwpu.edu.in

S. Kinger
e-mail: shakti.kinger@mitwpu.edu.in

I. Shanbhag
e-mail: ishaanshanbhag28@gmail.com

M. Ragib
e-mail: ragibmufaddal@gmail.com

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_72

1 Introduction

In medical world, surgeon will want to ensure that you are physically and mentally prepared for surgery. Before surgery, a person will undergo a series of examinations and testing. Before your operation, many members of your surgical team may ask you the same questions. This is because your surgical team has to gather as much information as possible in order to provide you with the best possible results. The time leading up to the operation is known as pre-op. It literally means "before the procedure." You will meet with one of your physicians during this time. This might be your surgeon or your primary care physician: This examination should be completed within a month of surgery. This allows your doctors to address any medical issues you may be experiencing prior to surgery. You will be questioned about your health over the years during this consultation. This is known as "taking your medical history," and it will be followed by a physical examination by your doctor. If you go to your primary care doctor for a pre-op exam, ensure the records are sent to your hospital or surgeon. Before surgery, your surgeon may recommend that you undergo several testing. Certain tests are required of all surgery patients. Others are only performed if you are at risk for particular diseases.

Blood tests such as a complete blood count (CBC) and kidney, liver, and blood sugar tests are common tests that your surgeon may request if you haven't had them recently (Fig. 1).

- A chest X-ray will be taken to assess your lungs
- ECG (electrocardiogram) to monitor your heart
- Tests to examine the lining of your bowels or stomach, such as a colonoscopy or upper endoscopy
- Cardiac stress test or other heart tests
- Lung function tests
- Other tests may be requested by some physicians or surgeons. This is determined by the following factors:

  Your age and overall health

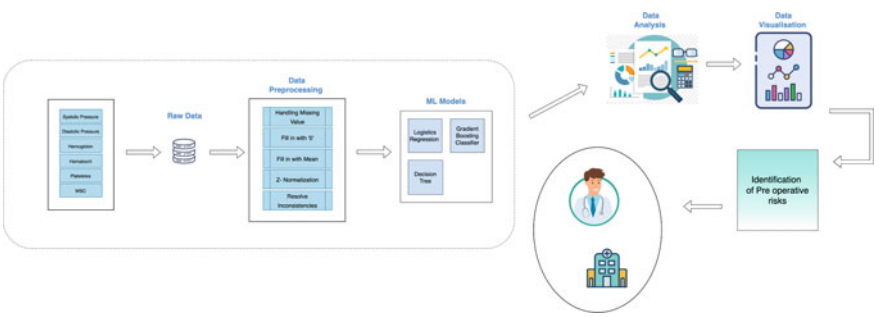


Fig. 1 Architecture diagram of proposed system

Any health concerns or issues you may have
The sort of operation you are having.

The architecture shows different components being used and how it interacts with the models. The raw data are collected in a single file, as we have to give this data as input to the models; preprocessing is done on the data which consist of different techniques depending on the dataset. The dataset required data cleaning and data transformation such as handling missing values appropriately, filling in the values, and normalization. The next is feeding the processed data to the models; now, the models have the algorithm to run on the dataset which gives accurate training to the models due to the processing done on the data. The results can be further analyzed to answer the unanswered questions. To make it appeal to the end user, a graphical representation of data visualization can be done on the result.

The following sections of the paper give us an idea about the previous work done by other authors in the same domain. Further, we have discussed processing the data and its data types. Later, we have explained the types of algorithms and their use in our proposed system.

2 Related Work

Preoperative evaluation has gotten a lot of attention in recent years. Many authors have used different sorts of data mining approaches. The authors proved that their study, which includes multiple datasets and algorithms for future research, gave a sufficient approach and confidence in detecting comorbidities that might lead to patient difficulties during anesthesia, surgery, or recovery.

The purpose of this study is to analyze various data mining procedures and statistical qualities that have been popular in recent years for interpretation of preoperative assessment with confidence by various authors in order to achieve diverse outcomes and approaches.

Several studies in the literature deal with identifying comorbidities that may contribute to patient difficulties during the anesthetic, surgical, or postoperative phase, as well as patient medical data such as blood pressure, sugar level, and creatinine. Some research, for example, employs machine learning classification and prediction models, while others employ deep neural network models.

Levett et al. showed in their research, how to identify patients who are at risk of a poor result and could benefit from preoperative intervention, accurate preoperative risk stratification is required. The causes of a patient's postoperative functional impairment are several. Both fixed (e.g., age, chronic comorbidities) and modifiable (e.g., smoking) risk factors can contribute to a bad result (e.g., physical activity levels, smoking). Preoperative functional capacity appears to be a good predictor of postoperative outcome, according to growing research [1].

Poldermans et al. have researched about an heart disease which is becoming a great complication among patients nowadays. The common cause of such complication

would be myocardial infraction. The inflammatory response can be reduced by statins and aspirin which in turn focuses on plaque stabilization. The mismatch between oxygen supply and demand is rectified by beta-blockers. This why preoperative risk assessment is excellent time for changing one's life habits and taking necessary precautions and receiving treatment in order to reduce life endangering cardiac risks [2].

Chassot et al. have identified that preoperative assessment of patients with risk of coronary artery disease has been the topic of several investigations. This paper proposes an new approach which can in turn help intermediate risk individuals to live longer and increase their long-term survival [3].

Healey et al. have found in this study that patients which had vascular surgical services were the ones which experienced different kinds of complications, and worst outcome was going to be death. The rates of complications for different kinds of patients were compared to relevant criterion norms [4].

Hollinger et al. had researched about how patients were screened prospectively for the first day and the tenth day post-surgery or at the time they were discharged from the hospital. The perioperative data were collected by them. It was observed that 11.5% of patients had delirium. So, it was discovered that delirium is common post-heart surgery and depends upon length of the time which the patients have spent of mechanical ventilation [5].

Charbuty et al. investigated a decision tree classifier known for its improved view of performance results. This paper outlines the latest research done in many areas, including medical disease analysis, text classification, user smartphone, and image classification. In addition, the details and dataset authors used in the techniques/algorithms used, and the results obtained in terms of accuracy are used in the summarized decision tree. Finally, when using a machine learning repository as a dataset, the highest accuracy achieved by the decision tree algorithm is 99.93% [6].

Evgeniou et al. study's findings were summed up in a report. Statistical learning theory is used to motivate SVMs. Using bounds on their capacity to predict future data, the theory characterizes the performance of learning machines. A constrained quadratic optimization problem is used to train SVMs [7].

3 Methodology

3.1 Dataset and Features

The dataset for this research study has been taken from Dr. B. R. Ambedkar Institute Rotary Cancer Hospital. The data were unrefined at first. It has null values. So, we normalized the data first. Then, we used the data for training the models. We implemented different algorithms, tested the accuracy of each algorithm, and optimized it. There were 13 parameters in the dataset.

This dataset will aid the model in determining whether or not the patient is at risk. 'Name', 'WBC', 'SGOT', 'SGPT', and 'Platelets' are some of the parameters in the dataset. The model can notify the patient to go to the doctor by establishing the range and verifying if the patient is inside it. A column in the dataset contains numerous values. We divided the multi-valued columns into independent columns to do proper preprocessing and offer accurate input to the model. For example, before preprocessing, white blood cells (WBC) and platelets were in the same column (Table 1).

For the values, there is a specific range of normal values on which the system will work upon. A practicing doctor checked these values. By these values, this table was built just for reference (Table 2).

Table 1 Data types of preoperative assessment variables

Preoperative variables		Total records	Count of missingdata	Data types
Name	Name	95	0	Nominal
Age	Age	95	4	Continuous
Sex	Sex	95	0	Categorical
Blood pressure	Systolic pressure	95	13	Continuous
	Diastolic pressure	95	13	Continuous
Hb/Hct	Hb-Hemoglobin	95	4	Continuous
	Hct-Hematocrit	95	44	Continuous
WBC/Plateletes	WBC	95	4	Continuous
	Plateletes	95	6	Continuous
Sugar F/PP/RN	F-Fasting	95	34	Continuous
	PP-Postprandial	95	86	Continuous
	R-Random	95	95	Continuous
Urea/Creatinine	Urea	95	4	Continuous
	Creatinine	95	6	Continuous
Na/K/Ca	Na-Sodium	95	7	Continuous
	K-Potassium	95	14	Continuous
	Ca-Calcium	95	55	Continuous
T Proteins/A/G	Total Proteins	95	14	Continuous
	A-Albumin	95	25	Continuous
	G-Globulins	95	76	Continuous
Bilirubin/D/I	Total Bilirubin	95	3	Continuous
	Direct Bilirubin	95	39	Continuous
	Indirect Bilirubin	95	51	Continuous
SGOT/SGPT	SGOT	95	10	Continuous
	SGPT	95	10	Continuous
ALKP04ASE	ALKP04ASE	95	14	Continuous

Table 2 Blood test data reference value

Blood test data for hemolytic jaundice		
BP	Systolic pressure	Diastolic pressure
	90–125	55–90
HB	6.5–10.5	
HTC	26–30	
WBC	4–11 (10^9/l)	
Platelets	180–500 (10^9/l)	
Creatinine	67–127 (ml/minute)	
T Protein	–	
Bilirubin/d/l	–	
SGOT	–	

4 Learning Algorithms

There are different machine learning algorithms which are used for prediction. We have used a few of them for our system in order to predict whether the patient is at risk to undergo a surgery.

The machine learning algorithms are primarily categorized into two types of algorithms which are supervised and unsupervised type of algorithms [8].

Supervised Algorithm: This is a type of machine learning algorithms which follows a training dataset. It mainly deals with labeled data to predict the output. It maps the function by using an input variable X to output variable Y.

Some of the examples of supervised algorithms are as follows: logistic regression, decision tree, support vector machine, etc.

Unsupervised Algorithm: This is a type of machine learning algorithm in which models are not supervised using a training dataset. It allows the model to work on its own, to discover different types of patterns which were previously undetectable. So, all the unlabeled data fall under this type of category. It allows users to perform difficult processing tasks as compared to supervised learning algorithm.

Some of the examples of unsupervised algorithms are as follows: K-means clustering, Apriori algorithm, neural networks, etc.

The primary goal of our study is to predict risk with a well-developed machine learning model. Our input data majorly are of a continuous data type. As a result, to predict a continuous variable with N inputs, models of regression and classification are applied. A discrete output is produced by using logistic regression to construct a relationship between a numbers of different input variables that are all independent. We have used 0.25 for testing and 0.75 for training with a random state as 1 for making the model. Furthermore, a decision tree is of both classification and regression techniques, as the name suggests it is a tree-structured classifier. The test set 0.34 is larger than the previous model; further, we have used entropy for measuring the

Table 3 Results of our proposed system

Algorithms	Hyperparameters	Training (%)	Testing (%)	Overall score (%)
Gradient boosting classifier	Learning rate/maximum depth/estimators	80	20	73
Decision tree	Entropy/random_state/min_samples_leaf	66	34	81.81
Logistic regression	NULL	75	25	83.34
Support vector machine	Linear	79	21	80

quality of split for decision-making. A support vector machine (SVM) is also used for both classification and regression problems. The dataset can be classified into different clusters or groups by using a single straight line; thus, a linear SVM model is a good fit for the problem. Gradient boosting classifiers are a collection of many weak learning models to generate a powerful prediction model. We have used leaning rate 1 and test set 0.2.

5 Results

A total of 94 patients (mean age 45 years;) blood test reports were included in this study. We performed the learning algorithms like support vector machine, linear regression, logistic regression, decision tree, gradient boosting regressor, and gradient boosting classifier, and we got the results as follows (Table 3):

6 Conclusion and Future Work

In this research paper, we found that the proposed machine learning framework for predicting complications for patient can be used in real-life decision support system which will help practitioners in planning and managing their resources wisely. In future, more models can be implemented to improve the accuracy, and extra preprocessing can be involved in order to make data suitable for all the models or even for a particular model which would aid in improving prediction accuracy.

References

1. Levett D, Edwards M, Grocott M, Mythen M (2016) Preparing the patient for surgery to improve outcomes. Best Practice Res Clin Anaesthesiol 30(2): 145–157. Available: https://doi.org/10.1016/j.bpa.2016.04.002. LNCS Homepage, http://www.springer.com/lncs, Accessed 21 Nov 2016
2. Poldermans D, Hoeks S, Feringa H (2008) Pre-operative risk assessment and risk reduction before surgery. J Am Coll Cardiol 51(20):1913–1924. https://doi.org/10.1016/j.jacc.2008.03.005
3. Chassot P, Delabays A, Spahn D (2002) Preoperative evaluation of patients with, or at risk of, coronary artery disease undergoing non-cardiac surgery. Br J Anaesth 89(5):747–759. https://doi.org/10.1093/bja/89.5.747
4. Healey M (2002) Complications in surgical patients. Arch Surg 137(5):611–618. https://doi.org/10.1001/archsurg.137.5.611
5. Hollinger A, Siegemund M, Goettel N, Steiner L (2015) Postoperative delirium in cardiac surgery: an unavoidable menace? J Cardiothorac Vasc Anesth 29(6):1677–1687. https://doi.org/10.1053/j.jvca.2014.08.021
6. Charbuty B, Abdulazeez A (2021) Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends 2(01):20–28. Available: https://doi.org/10.38094/jastt20165
7. Evgeniou T, Pontil M (2001) Support vector machines: theory and applications. In: Machine learning and its applications, pp 249–257. Available: https://doi.org/10.1007/3-540-44673-7_12. Accessed 25 April 2022
8. Han J, Kamber M, Pei J (2012) Data mining. Elsevier, Amsterdam

A Comparative Analysis of Various Techniques of Data Leakage Detection in Different Domains

Kiran Patil, Harsha Sonune, Soniya Devikar, Vrushali Chaudhari, and Isha Ayachit

Abstract With the steep growth in information technology and its global reach, as well as the common citizen's ever-increasing reliance on technology, data privacy and security have become a major source of concern for individuals all over the world. In today's era, computing devices like virtual servers, databases, physical servers, databases, and many more devices are occupied with confidential data. This paper is an exploratory case study that analyzes the various algorithms and methods proposed across the various domains, and a comparative analysis was done.

Keywords Data leakage detection · Android · Networking · Cloud computing · Machine learning · Guilty agent · Privacy · Watermark · Fake object · Bigraph

1 Introduction

Data is critical and an important aspect as it contains sensitive information related to user privacy, finance, social privacy, etc. This data is a precious asset to companies and organizations for project and business insights. Nowadays, almost all of our data is becoming digital, and because of online sharing platforms, data privacy stands a major risk.

In Jan 2021, about 35 million user accounts data from Juspay was for sale on dark Web. An un-recycled access key was used to steal details of around 35 million

K. Patil (✉) · H. Sonune · S. Devikar · V. Chaudhari · I. Ayachit
MKSSS's Cummins College of Engineering for Women, Pune, Maharashtra, India
e-mail: kiran.m.patil@cumminscollege.in

H. Sonune
e-mail: harsha.khedkar@cumminscollege.in

S. Devikar
e-mail: soniya.devikar@cumminscollege.in

V. Chaudhari
e-mail: vrushali.chaudhari@cumminscollege.in

I. Ayachit
e-mail: isha.ayachit@cumminscollege.in

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_73

consumer accounts, including disguised card data and fingerprints, from a server. In March 2021, around 45 lakh user data was leaked from known organization Air India, whereas Dominos reported 18 crore users' data breach in May 2021.

A single corporation may have access to the personal information of millions of customers, which it must keep private in order to protect consumers' identities. Malicious third parties or hackers can steal this information for their own gain at the expense of ordinary people and businesses' private information. Data can be exposed in a variety of ways. It could be the result of a cyberattack designed to steal data, or it could be an employee purposefully leaking or selling private data to a third party. Domains covered in this paper are Android, ML, deep learning, networking, cloud computing, and fake/guilty agents.

2 Literature Survey

There are many different strategies or techniques that are developed or suggested so far for detection and prevention of data leakage. Here are some of the strategies/techniques used:

Cam recommended uitXROM for the purpose of detecting leakage of sensitive data in custom created Android firmware by inspecting the connection between apps that were pre-installed [1]. This system consists of three major modules: APK extractor (extracts pre-installed apps), APK analyzer (detects sensitive data flow via multi-applications), and path matcher (analyzes data paths based on path entry and path exit points). The results of the experiment show that the system can detect several applications that are already installed that leak sensitive data from over 280 ROMs that are found online or downloaded from the Internet.

The authors of the paper Naik and Gaonkar suggested watermarking to detect data leakage in clouds [2]. Watermarking approach takes data to be conveyed in picture, and a quick response code is inserted. The generated QR code watermark is embedded into cloud data. Current data characteristics are examined to estimate tampering in data, and the guilty agent is identified by taking out watermark and collating watermark data to the agent's details. Existing systems can provide security by encryption utilizing numerous techniques; however, the proposed paradigm provides both security and detection.

Ranchal and others proposed a security enforcement implementing security protocols in composite online services structure [3]. It makes use of active bundles. When the AB engine detects an assault, the EPICS technique enables dynamic data destruction to prevent disclosure, which is not possible with models that treat data as passive entities. The proposed framework is backwards compatible with existing service architecture and meets the real-time needs of Web service interactions.

Rastogi proposed a Uranine technology that monitors Android apps in real time to detect privacy breaches [4]. Upon receiving an app, it is converted by Uranine to a custom IR which can be implemented during runtime for taint propagation. The IR that has been instrumented is then translated to bytecode, following which a fresh

app is created and reassembled to create a new app that can be downloaded and installed on an Android device. Privacy leaks are automatically tracked when the instrumented application runs. Uranine implements Android apps which is why it does not require platform support for tracking information flow. Uranine has good accuracy and a low performance overhead when compared to other practices.

For detecting attacks in huge networks in real time, Wu and others present a method for converting raw traffic vector form to image data form [5]. The image form data has the potential to reduce the number of computing parameters. The study suggests a technique for using the training sample quantity as a basis for determining the weight coefficients of each class's cost function and the purpose of which is to increase the detection of accuracy of the unbalanced traffic dataset. Experiment results indicate that the suggested CNN intrusion detection model performs better than the pre-existing intrusion detection techniques.

Min and others propose TR-IDS, an unique intrusion detection framework that uses both features designed manually and payload features to increase performance [6]. To extract primary features from payloads, it uses two NLP techniques, namely word embedding and Text-CNN. Word embedding maintains semantic relationships among bytes while reducing feature magnitude, and Text-CNN is then utilized to pull out features from all payloads. For the final categorization, it used a complex random forest algorithm.

The technique introduced by Liu and team is a CNN-based multi-classification detection system for network intrusion [7]. The experimental findings were compared to deep learning models like DNN, GRU-RNN, LSTM-RNN, and others using the KDD-CUP99 and NSL-KDD datasets. The CNN-based model enhances recall and accuracy, minimizes false positive rate, and produces improved outcomes for unexpected attacks detection.

Alrawashdeh and Carla Purdy introduced a deep learning-based method for detecting anomalies that has a high accuracy and efficiency rate on the remaining 10% of the testing data from the KDDCUP99 dataset [8]. This method is based on a deep belief network fine-tuned with logistic regression softmax. To increase the overall performance of the network, a multiclass logistic regression layer was implemented and trained with 10 epochs on the upgraded pre-trained data. In addition, they simulated a network with a short training time and minimized the dataset's preprocessing.

Gupta and Singh have introduced a discriminatory criminal model that identifies malicious entities involved in a confidential data breach and provides security to prevent breach of personal data [9]. This requires addressing the issue of data breaches. Based on the assigned data to different clients or agents, this implementation foretells the guilty agent. In the model, the supplier distributes the given data item shared by numerous agents. Bigraph is used to represent these things. We use this graph to find the matrix, which gives us a list of the agents to whom each data object has been assigned. If an agent discloses sensitive data objects to a malicious third party and the distributor later discovers them in an illegal location, the distributor is liable.

Govinda and Divya Joseph have employed the data allocation methods to determine the exact moment and also the guilty agent. This demonstrates how the distributor can use a basic approach of introducing bogus items to prime line identification during the initial distribution phase [10]. These bogus objects that are injected have no relation to the actual data, yet they appeal to the distributed agent as real data. This notion is synonymous with the concept of embedding watermarks. Where there are parallels to be drawn based on the fact that object insertion works in a similar way to hiding a watermark. The distributor may readily identify the agent who is responsible for this with the help of these false things. This method also delivers evidential proofs that accurately identify the guilty agent.

The most famous prevention technique for data leakage in peer-to-peer networks is the data loss prevention technology also known as DLP, but it has some disadvantages. For example, due to the high-detection-error rate, it might prohibit transferring of a normal file externally from the organization, and it can also cause breach of privacy of internal staff for filtering of data. Hence, Chae proposed a system that can prevent internal data loss for employees and customers via a peer-to-peer network [11]. The suggested system can identify the personal data and expel it from the sharing file using the risk factor of privacy data leakage. This procedure also addressed DLP system issues like high-detection-error in data privacy and concerns of invasion of privacy.

To prevent data leakage by internal staff, companies enter the behavior pattern associated with data leaking into the system ahead of time and identify the employee as the staff who released the data, whose behavior pattern is then identified. However, because the data was not entered into the system, the reason for the data leakage cannot be correctly determined if it is released according to the pattern of security log occurrence. Seo and Myung-Ho Kim propose a system that uses convolutional neural networks to identify the cause of data leakage, in which the data leakage pattern is defined as a series of security logs that can appear immediately at the time of data leakage rather than being fed into the system [12]. As a data leakage judgment scenario, these security logs might arise as a result of an association analysis algorithm, i.e., the Apriori algorithm.

The majority of data leakage events occur while employees are performing routine operations, such as sending email that contains sensitive information accidentally. Sensitive information, such as bank records and credit card scores, is sent through email. Wang suggested an email protection system by explaining threats to an email and why it is important to protect emails [13]. The key advantage of building this system on gateway is that it safeguards the company's important information from rivals.

A research that analyzes the present condition of security monitoring managed by three Korean enterprises and suggests certain risk scenarios about unauthorized access is assessing risk scenarios regarding security monitoring system vulnerabilities with an emphasis on information leakage by insiders. The study has gathered each business's policy on security, systems for security monitoring and the utilized system log for the case analysis. As a result of this research, four risk situations that

are probable to occur in future were discovered, as well as threats that were hard to identify using the current security monitoring system [14].

Time stamps are very important in data leakage prevention for providing permission to access a certain data because the data is confidential during a specific period of time [15]. Because the same data could become non-confidential after the timestamp, a technique for preventing data leakage with time stamps was proposed.

Karaçay and others proposed a protocol for detecting intruders over encrypted Internet data in which the detection model and the entire network are kept private [16]. The authors used a homomorphic encryption algorithm as the attack structure and used a personally evaluated decision tree on the network data. The analysis demonstrates that by regulating a number of parameters such as the list of laws, the numeral characteristics, and the list of categories in feature representation, security, and privacy can be improved. The proposed method improves privacy as well as quality in terms of execution time.

Another method of detecting data leakage is to use a solution that can be delegated and deployed in a reasonably ethical detection environment. Acts to detect data leakage technique named "fuzzy fingerprint" is used to improve data privacy. This procedure is based on a one-way calculation of sensitive data. This enables data owners to securely handle content monitoring responsibilities to DLD providers while keeping important documents protected [17].

3 Observation

Data leakage detection system can be covered under various domains. These domains are guilty/fake agent, networking, cloud computing, Android, machine learning, artificial intelligence, and deep learning.

The guilty/fake agent domain introduces techniques for detecting data leaks produced by malicious entities intentionally or unintentionally transmitting confidential data to unauthorized third parties.

Networking domain includes P2P file sharing and mobile ad hoc network. Using the P2P network, which is fundamentally a direct link between computers. P2P file sharing allows registered users to exchange and share files with fellow peers without the usage of networks. The disadvantage is that it may allow internal personnel and customers' personal information to be leaked.

Cloud computing includes techniques such as watermarking, use of active bundles, and other methods to detect data leakages in cloud and other Web services.

Android mainly focuses on analyzing relationships of pre-installed applications and technologies that monitor Android apps in real time to detect privacy breaches.

Machine learning, artificial intelligence, and deep learning mostly are about intrusion detection systems. This domain does not cover much about data leakage detection but is useful for predicting various kind of attacks like denial of service, remote to local, user to root, and probing for follow-up procedures.

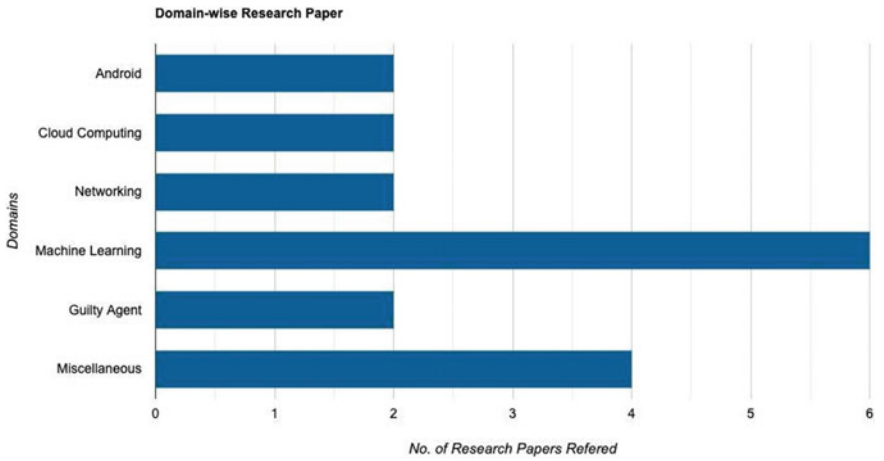


Fig. 1 Domain-wise comparative analysis of research papers taken for reference

Figure 1 represents the following statistics of research papers from various domains:

Android—2
Cloud computing—2
Networking—2
Machine learning—6
Guilty agent—2
Miscellaneous—4.

4 Conclusion

There are several ways for leakage of data to happen, and the research papers cover various techniques and methods to detect the data leakage. Data can be leaked by an insider or an employee of the organization, or by a third party agent to whom the distributor shares data with, or by an external attack like hacking. These are some of the ways that are covered by this report as well as methods and techniques that can be implemented as solutions for these problems. According to the troubles faced by an organization or a company, appropriate methods can be integrated within their systems categorized by the various results and output of each technique.

Among the machine learning, artificial intelligence, and deep learning papers, we found that the information-retrieval-based technique is suitable for data leakage detection. This technique is useful for Web data leakage detection for sensitive data. The system monitors the Web and collects information about Web documents according to user's preferences. If a document on the Web appears to be semantically

similar to confidential user documents, the system alerts the user to the possibility of data leaking.

Out of all the techniques in the cloud computing domain, we found the use of active bundles for data leakage detection more suitable. This technique protects data from malicious cloud administrators by providing data integrity and confidentiality, and the method also allows for access control based on roles and attributes.

In the Android domain, uitXROM technique is more suitable for data leakage detection. This approach detects sensitive data leaks in already installed apps on modified Android ROMs.

In the field of networking, wireless networks like MANET are a very unique application due to scalability and mobility. Data leakage detection and reduction (DLDR) with lightweight cryptography is a method for detecting and reducing data leakage via the Internet. Use of lightweight cryptography and S-Max algorithm supports data confidentiality and reduces scope of data leakage.

There are many algorithms and methods proposed to find out the guilty or fake agent who maliciously leaks confidential information and based on the organization's needs and requirements as well as the based on the particular region which is in essential demand for security and privacy, the appropriate method can be utilized effectively from the comparative analysis provided in this analytical case study.

References

1. Cam NT (2017) Sensitive data leakage detection in pre-installed applications of custom Android firmware. https://doi.org/10.1109/MDM.2017.56
2. Naik R (2019) Data leakage detection in cloud using watermarking technique. In: 2019 international conference on computer communication and informatics, pp 1–6
3. Ranchal R (2019) EPICS: a framework for enforcing security policies in composite web services. https://doi.org/10.1109/TSC.2018.2797277
4. Rastogi V, Qu Z, Mcclurg J, Cao Y (2015) Uranine: real-time privacy leakage monitoring without system modification for android, vol 2, pp 256–276. https://doi.org/10.1007/978-3-319-28865-9
5. Wu K, Chen Z, Li W (2018) A novel intrusion detection model for a massive network using convolutional neural networks. IEEE Access 1. https://doi.org/10.1109/ACCESS.2018.2868993
6. Min E, Long J, Liu Q, Cui J, Chen W (2018) TR-IDS : anomaly-based intrusion detection through text-convolutional neural network and random forest, vol 2018
7. Liu G, Zhang J (2020) CNID: research of network intrusion detection based on convolutional neural network, vol 2020
8. Alrawashdeh K, Purdy C (2016) Toward an online anomaly intrusion detection system based on deep learning. https://doi.org/10.1109/ICMLA.2016.167
9. Gupta I, Singh AK (2017) A probability based model for data leakage detection using bigraph, pp 1–5
10. Govinda K (2017) Dynamic data leakage using guilty agent detection over cloud. In: 2017 international conference on intelligent sustainable systems (ICISS), pp 744–746
11. Chae C, Shin Y, Choi K, Kim K, Choi K (2015) A privacy data leakage prevention method in P2P networks. https://doi.org/10.1007/s12083-015-0371-x
12. Seo M (2017) An advanced data leakage detection system analyzing relations between data leak activity. Int J Appl Eng Res 12(21):11546–11554

13. Wang S (2017) Data leakage prevention: e-mail protection via gateway data leakage prevention: e-mail protection via gateway
14. Kim K (2019) A study on analyzing risk scenarios about vulnerabilities of security monitoring system: focused on information leakage by insider. Springer, Berlin
15. Peneti S, Rani BP (2016) Data leakage prevention system with time stamp. In: ICICES, pp 3–6
16. Karaçay L, Sava E (2019) Intrusion detection over encrypted network data. https://doi.org/10.1093/comjnl/bxz111
17. Shu X, Yao D, Bertino E (2015) Privacy-preserving detection of sensitive data exposure. IEEE Trans Inf Forensics Secur 10(5):1092–1103. https://doi.org/10.1109/TIFS.2015.2398363

Optimized Closest Pair Computation with CPU-GPU Combined Model

Prafullata Auradkar, G. R. Gagan, Sarthak Deva, Navya Eedula, Mrudhulraj Natarajan, Subramaniam Kalambur, and Dinkar Sitaram

Abstract Spatial data processing had been a dominant contributor in a wide assortment of applications including health care, urban planning and infrastructure designing. As the complexity of data has increased due to higher dimensions, diverse approaches have been successfully tested for nearest neighbor queries, clustering algorithms, etc. These algorithms have been optimized and adapted for CPU, GPU and hybrid models as combination of CPU and GPU. Closest pair (CP) computation is a frequently used operation in range and distance-based queries. Although there are different algorithms and optimizations available for CP computations, suitability of GPU for this operation is yet to be explored. Hereby in this paper we propose CPU-GPU Hybrid Model, to optimize the closest pair (CP) of points problem concerning two-dimensional floating point values. As part of the combined model, the sorting phase is implemented on CPU and computation of closest pair is implemented on GPU. In this work, three CP algorithms are analyzed on CPU and CPU-GPU combinations. With the proposed algorithm on the CPU-GPU combined model, up to 12 times speedup is achieved in the closest pair computation time. Several parameters were tuned to work with the device-specific features and the overhead in terms of data transfer is also analyzed.

Keywords Closest pair · CPU-GPU combined model · Divide and Conquer · Quick CP

1 Introduction

Computational geometry and spatial operations have a large multitude of applications in health care, navigation systems, airspace design, air traffic control and urban planning and smart cities, etc. [1, 2]. These operations encompasses computations for closest pairs that would assist in range and distance-based queries [3, 4]. Additionally, the closest pair of points computations has applications in pattern mining

P. Auradkar (✉) · G. R. Gagan · S. Deva · N. Eedula · M. Natarajan · S. Kalambur · D. Sitaram
Department of Computer Science and Engineering, PES University, Bangalore, Karnataka, India
e-mail: prafullatak@pes.edu

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_74

for closest pair of sub-sequences [5] and finding closest l-mers which is essentially pattern mining in biological data [6]. Published work suggests optimizations in terms of maintaining closest pair when data set is dynamic [7]. Authors in [8] suggest a progressive approach in computing the closest pairs by considering partial solutions rather than the exact solutions for the closest pair with an acceptable error threshold as defined for the application considered. Computational complexity with this approximation is claimed to be $O(n\log(n))$.

The existing algorithms for the closest pair of points computation include the brute force method with an upper-bound of $O(n^2)$ time complexity. A further improvement introduced is the well-known divide and conquer approach bound at $O(n(\log(n))^2)$ time complexity. The Quick CP algorithm [9] though has a worst case complexity of $O(n^2)$,introduces a minor yet effective change to the brute force algorithm which reproduces an average time complexity of $O(n\log(n))$. It relies on the observation that two points sorted in x-axis and/or y-axis will reflect the closest pair of points.

The parallel processing power of GPU is used for various spatial operations but has not yet been explored for the closest pair problem.

Our Contributions: Proposed algorithm for closest pair computation exploits the best behavior of CPU and GPU.

- CPU-GPU combined approach to compute closest pair efficiently.
- Both local and global closest pairs are computed over GPU cores.
- Based on the point data set size, a balance between the number of batches and batch size is computed.
- Sorting phase in all three algorithms is implemented with the CPU. Overall performance gain observed is due to the optimized computation time.

2 Related Work

The divide and conquer is a well-known approach that breaks down the given problem recursively into sub-problems and consolidates individual observations to achieve the result. This method employed on the closest pair problem retrieves the result within $O(7n\log(n))$ time complexity with at most $O(7n\log(n))$ euclidean distances and $O(n)$ square roots computed [10]. An improvement on the former method is to sort the points along the y-axis in $O(n\log(n))$ time, followed by closest pair computed in $O(n)$ time complexity [11].

Approach proposed in [12] proposes a lemma according to which it is necessary and sufficient to check only the subsequent three points among the y-sorted points to detect overlapping pair of points and hence find the closest pair of points in a given dataset. Ovidiu Daescu and Ka Yaw Teo further proceed to prove the lemme by deriving a tight geometric bound for a specific step in the combine phase of the well-known divide and conquer algorithm. The experimental results for this algorithm prove to be better than the classic divide and conquer algorithm due to the reduction

in the number of computations. The paper further implies that the lemma can be used in the future to create an adaptive heuristic approach for various distributions of a 2D point dataset.

Survey of GPU-based sorting algorithms is presented in [13]. It was indicated that CUDA-Quick Sort has shown better parallelization than other algorithms. Authors have also discussed the desirability of GPUs over multi-core CPUs though both processors are used to achieve performance enhancements via parallelism. CUDA-quicksort in detail can be found in [14].

In [15] an external sorting algorithm was proposed where data transfer and GPU computations were handled in an overlapped manner that masks the transfer time overhead and multiple GPU utilization was proposed with better load balancing. According to [14], it has been proposed that that inter-block communications can be enabled to improve the GPU memory access for sorting.

There are several approaches proposed for closest pair computations that employs CPUs and/or GPUs. Eldawy, Li, Mokbel and Janardan in [16] have commissioned multiple cores of CPU through Map-Reduce Programming for Computational Geometry operations. Closest pair operations are implemented as parallel computations where local closest pairs are computed for the spatially partitioned subsets of points in Map function and subsequently a global closest pair is computed in the reduce function.

The incorporation of Geometric predicates [17] and geometric models [18] in Compute Unified Device Architecture (CUDA) libraries extended the ability of GPUs to apply clustering algorithms [19] as stand alone or as a hybrid model of CPU and GPU combinations [20]. But closest pair computations using GPU or a combined model with CPU and GPU are not yet explored. This empirical study brings out the opportunities and analyzes the limitations in the proposed algorithm with suitable test cases. A Hybrid architecture of CPU and GPU is proposed in [21] to train node embeddings. It explores the effectiveness of combined model for computation(CPU and GPU) that have been endeavored in several computational tasks. It suggests parallel approaches where CPU cores are utilized to generate training data for the network and multiple GPUs are utilized for training the node embeddings. A novel approach was proposed that masks the data transfer latency and synchronization needs. CUDA architecture implements a Single Instruction Multiple Thread (SIMT) architecture which provisions greater performance and programmability on graphic devices [22]. The proposal made by authors in [22] for a heterogeneous framework, effectively utilizes the host CPU resources and CUDA kernels of the GPU to enhance performance. A framework is implemented to distribute workload to enable concurrent running of CUDA kernels on CPU and GPU in an optimized ratio. This resulted in empirical values very close to the theoretical analysis. This concludes that CPU cores can successfully be utilized for CUDA kernels. CUDA Library in programming, provides optimized functions to initiate kernel codes that are used to achieve data parallelism. Each kernel describes the work of a single thread which can be invoked on several thousands of threads in parallel. High degrees of parallelism can be achieved by extensive multi-threading and utilization of maximum number of cores. This hides the global memory latency shown by GPUs. However, finding

the right balance between the amount of computation happening per thread and the number of parallel threads being invoked can be a daunting task [23].

Every GPU possess separate local memory buffers which require explicit data transfers from CPU to GPU. The limited device memory in the GPU brings about the need to partition the data and iteratively transfer it from the host memory. A solution proposed in [24] to overcome this problem is the concept of double buffering which ensures concurrent GPU processing with data transfers using twice as many GPU threads.

Work in [25] suggests a variation of the standard divide and conquer approach by dividing the problem into n different spaces instead of 2. This division of the problem into n sub-problems is done recursively. Each sub-problem has equal number of points within which computation of closest pair is no longer a recursive process. Authors in [26] have proposed an interesting way of reducing the number of computations required to find the closest pair of points by employing dynamic warping. The distances between the set of points and an initially computed reference point is calculated. The number of comparisons were brought down from 2.4×10^{11} of the divide an conquer algorithm to a mere 6000 for a point data set of size 10 million.

3 Computing Closest Pair

3.1 Divide and Conquer Approach

This is a classical approach to find the closest pair in $O(n\log(n))$ time complexity on the CPU. The data set of two-dimensional points is sorted according to the x-axis. The mid-point along the x-axis is used to divide the data set into two equal halves after which a left sub-problem and a right sub-problem is obtained. The recursion occurs till the smallest sub-problem consists of three points is obtained whose solution is obtained by brute force method. After the computation of the minimum distance in these sub-problems, a check is made to ensure that there is no pair of points exists in the strip of length d along the vertical line between the two sub-problems that is lesser than the previously computed smallest distance. If so, this new set of points are taken as the closest pair.

3.2 Quick CP Algorithm

Paper [5] proposes the Quick CP algorithm which is based on the brute force approach by capturing the closest pair of points w.r.t the sorted projections of all the points. Although theoretically, the worst case time complexity of this algorithm is $O(n^2)$, it has been proven to be faster than the classic divide and conquer algorithm with an average time complexity of $O(n\log(n))$. This optimization is due to the the break

condition, which is used to compare if the current closest distance in the loop is lesser than the computed closest distance. If this is not the case, then the remaining part of the loop is skipped. This small yet drastically effective change is well suited for CPU implementation.

3.3 Proposed Methodology: Hybrid CP Algorithm

Let P be the point data set which consists of points of the form $P = \{P_1, P_2, \ldots P_i \ldots, P_n\}$ such that P_i is a two-dimensional point of the form (x_i, y_i). All the points are sorted using the standard **qsort** function (from the C library) on the CPU according to their x coordinates.

As the GPU Kepler architecture used in this study supports only 16 GB of points being present in the memory of the GPU at a time, the sorting steps for X-coordinates and later for Y-coordinates are accomplished on CPU.

The batch size value indicates the number of points in every thread. Let this value be BS. This heuristic value is a trade off between batch size and the number of batches. Increasing the batch size would increase the computation in each thread while increasing the number of batches would decrease the number of points which would essentially increase the communication cost.

Hence, identifying the optimum batch size before initiating the kernel computations is necessary. Empirically the batch size is varied for different input file sizes and the average execution time was plotted against the batch size The bell-curve as seen in Fig. 6 suggests an optimum batch size of BS = 4. Hybrid CP approach as proposed by this paper is illustrated in Fig. 1.

If the number of points in P is lesser than $BS * TotalNoOfThreads$ then the batch size is set as BS. In this case, the number of threads that will perform the computation will be $\lceil (n/batchSize) \rceil$ and the number of blocks that will be $\lceil (usedThreads/maxNoTBP) \rceil$.

Else, if the number of points in P exceeds this threshold, then the maximum number of threads ($maxNoOfThreads$) and blocks ($maxNoOfBlocks$)available are used for the computation. The batch size is changed to to accommodate all points with the available threads to $\lceil (n/usedThreads) \rceil$.

Each batch is further sorted with respect to the y-coordinate on the CPU using the same C-function **qsort**.

A kernel call, $BlockFunc$ is made to compute the closest distance between the points in each batch. These distances are stored in a temporary array $localMins$. $minDistance$ stores the minimum of all the values in $localMins$.

Another check needs to be performed in the case that there exists a minimum distance across the batches which is lesser than the computed $minDistance$. This is done via the kernel call, $StripFunc$ and stored in the temporary array $localSplitMins$.

Algorithm 1 Closest Pair Algorithm

BS // Heuristic Value for batch size
$Px \leftarrow sortByX(P)$
if $0 < n < (BS * totalNoOfThreads)$ **then**
  $batchSize \leftarrow BS$
  $usedThreads \leftarrow \lceil \frac{n}{batchSize} \rceil$

  $usedBlocks \leftarrow \lceil \frac{usedThreads}{maxNoTBP} \rceil$
else
  $usedThreads \leftarrow maxNoOfThreads$
  $usedBlocks \leftarrow maxNoOfBlocks$
  $batchSize \leftarrow \lceil \frac{n}{usedThreads} \rceil$
end if
$Py \leftarrow sortBatchByX(Py)$
$localMins \leftarrow [d_1, d_2, d_3, \ldots, d_n]$ $\quad //n = usedThreads$

BlockFunc$<<< usedBlocks, maxNoTBP >>> (n, Px, Py, batchSize, localMins)$
$minDistance \leftarrow minimum(localMins)$
$localSplitMins \leftarrow [s_1, s_2, s_3, \ldots, s_n - 1]$
  $//n = usedThreads, s_1 = s_2 = \ldots = s_(n-1) = minDistance$

StripFunc$<< usedBlocks, MaxNoTBP >> (n, Px, Py, batchSize, localSplitMins)$
finalMinDistance $\leftarrow minimum(localSplitMins)$

Algorithm 2 BlockFunc Definition

$start \leftarrow gridId * batchSize$
$end \leftarrow minimum(start + batchSize, n) - 1$
$minD \leftarrow QCP(Px[start : end], Py[start : end])$

Algorithm 3 StripFunc Definition

$start \leftarrow gridId * batchSize$
$end \leftarrow minimum(start + batchSize * 2, n) - 1$
$midPoint \leftarrow \lceil \frac{start+end}{2} \rceil$
$midVal \leftarrow Px[i].x - localSplitMins[gridId]$
for i=$midPoint$ to $start$ **do**
  if $midVal - Px[i].x > localSplitMins[gridId]$ **then**
    BREAK
  end if
end for
for $j = midPoint$ to end **do**
  if $midVal - Px[j].x > localSplitMins[gridId]$ **then**
    BREAK
  end if
end for
for $k = i$ to j **do**
  for $l = k$+1 to j **do**
    $localSplitMins[gridID]$ =
    minimum($localSplitMins[gridID]$,
    dist($Px[k]$,$Px[l]$))
  end for
end for

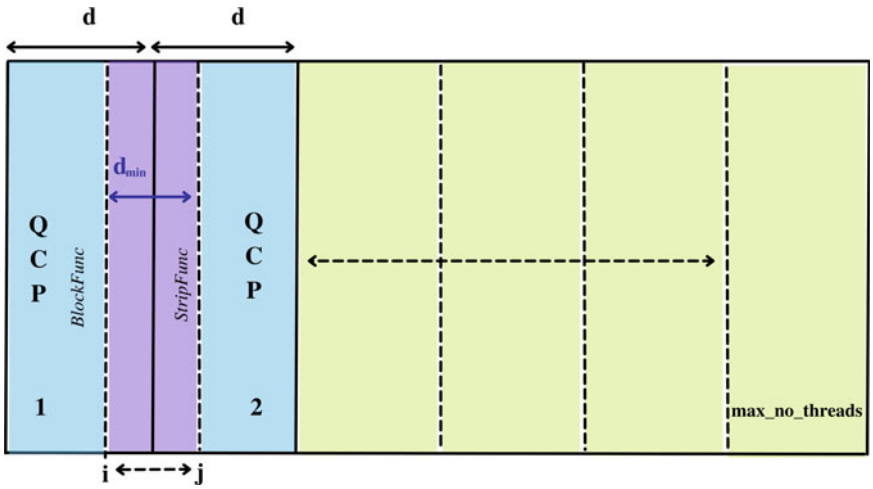


Fig. 1 Block diagram depicting the Hybrid CP approach in one block

The final closest distance, $finalMinDistance$ is hence the minimum of all values in $localSplitMins$ and the pair of points in P corresponding to this value is the solution.

The following is a two-dimensional diagrammatic representation of the algorithm employed:

4 Results and Discussion

4.1 Experimental Setup

The GPU used is Intel Skylake C s, Quadro RTX 5000, GT 710 which has Kepler architecture with a CUDA framework version 11.0. The GPU consists of 3072 cores and its configuration is such that it allows a maximum of **65,536 blocks** and **1024 threads** per block allowing the total number of threads to be 67,108,864. **Profilers-** NVPROF profiler was used to track the timings of kernel calls, API calls, memory transfers and for relative time spent in CPU versus GPU. **Datasets-** The data sets are generated using Spatial Hadoop generator. These are unique points generated with random distribution within a given minimum bounding box (MBR) with the file sizes ranging between **50 MB** to **15 GB**. The data sets were generated in spatial partitions of size **512 MB** which were subsequently merged before the data was sent to the CPU for processing. Table 1 illustrates the number of points present in these test data sets along with file sizes. Each point in this synthetic data set is represented with floating point values of up to 8 decimal places for the two-dimensional coordinates. While geographic locations are represented as points with longitude and latitude as

Table 1 Different file sizes and corresponding number of 2D floating point values

File sizes (in GB)	Number of points
1	29,545,119
5	147,726,372
10	295,453,891
12	354,543,625
15	443,179,083

floating point coordinates, these synthetic point data come closer in terms of the computational complexity.

Determining the number of batches, reading the point data set from the CPU memory and subsequent sorting of on X-axis and Y-axis is performed on the CPU using standard C library functions. GPU threads are then used to find the local closest pair within each batch. Then, the partial solutions are reprocessed with a second phase of GPU kernel call to compute the global closest pair. Unlike the Divide and Conquer approach the Hybrid CP approach does not utilize recursion to solve the sub-problems. It is empirically observed that non-recursive algorithms show better performance than recursive ones [27]. This led to a significant improvement in computation time. As proposed in [23], it is crucial to ascertain an ideal batch size for the the GPU threads so that the balance between the data transfer overhead and computation is maintained. Discussion in [12], indicates that the comparison of at most 3 subsequent points is necessary and sufficient to compute the closest pairs. As illustrated in Fig. 6, different batch sizes are tested to validate the ideal time taken to complete the job. It is interesting to note that the ideal batch size happens to be 4, which is irrespective of the input file size tested. The values used to plot this graph are the average across all the different input file sizes.

As indicated in the proposed algorithm, in the case that the file exceeds $batchsize * TotalNumberofthreads$, the batch size is increased. However, as the test cases in our study are limited to 15GB, impact of the alternative batch sizes were not explored. The calculation of the batch size, reading the point data set and subsequent sorting of the point data in the two-dimensional coordinate space, is performed on the CPU. Once the batches are formed, GPU threads are initiated to compute the closest pairs per batch. This is followed by another GPU Kernel function to compute the global closest pair. The three algorithms that are discussed in Sect. 3 are Divide and Conquer (DAC), Quick CP (QCP) and Hybrid CP(HCP). DAC and QCP are implemented on CPU. All the three algorithms have a sorting phase which is implemented on CPU. Figure 2 illustrates the closest pair computation time taken by all three algorithms. As it can be seen, there is up to a 12.18 times speedup achieved in the computation time using HCP. Table 2 cataloged the computation times exhibited by all the three algorithms. This was excluding the sorting phase and also the data transfer phase. This improved computation time accounted for 91.8% decrease in the overall completion time of HCP as illustrated in Fig. 3. Due to the data transfer overhead the

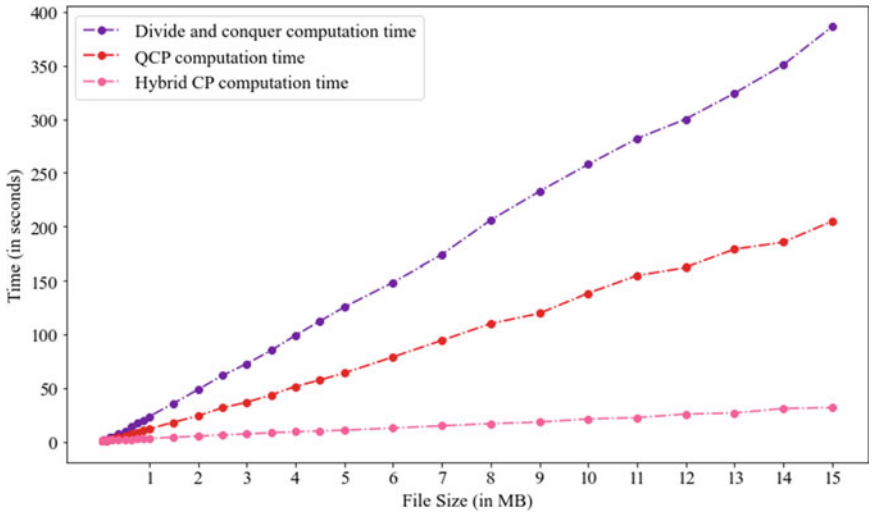


Fig. 2 Comparison of computation time of the three algorithms

Table 2 Computation time (in seconds) by the three different algorithms

Algorithm employed	5 GB	10 GB	12 GB	15 GB
Divide and Conquer on CPU	125.320	257.734	299.960	386.015
QCP on CPU	63.953	138.291	162.080	204.748
Hybrid-CP on CPU and GPU	10.647	21.151	25.740	31.672

speedup seen in the computations part is not as it is translated to the overall completion of the closest pair. The impact of these phases on the overall completion time is demonstrated in Figs. 4 and 5 (Fig. 6).

It would be relevant to explore the efficiency of HCP algorithm in comparison to a multi-core implementation of CP computation as seen through Map-Reduce computation of [16]. The Map-Reduce programming model parses data in parallel using the Mappers on HDFS data, followed by the computation of the closest pair of points across all mappers and finally determine the global minimum using the reducer. So to avoid this discrepancy in the data transfer overhead, only one test case with size 512 MB is cross-validated with HCP on GPU. It is observed that Map-Reduce took 56.24 s while HCP on GPU completed in 17 s. Although this result illustrates a speedup, it can not be packaged as a overall speedup for HCP, due to the corresponding data transfer latency. The possibility of testing larger data sets on HDFS is not there as computation happens across multiple machines in Hadoop while it happens across multiple threads in a GPU architecture. The data being distributed

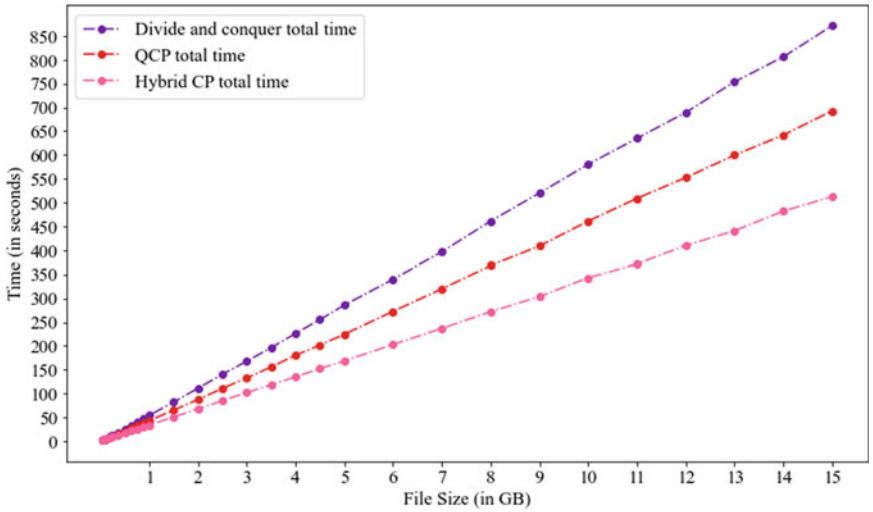


Fig. 3 Comparison of total time of the three algorithms

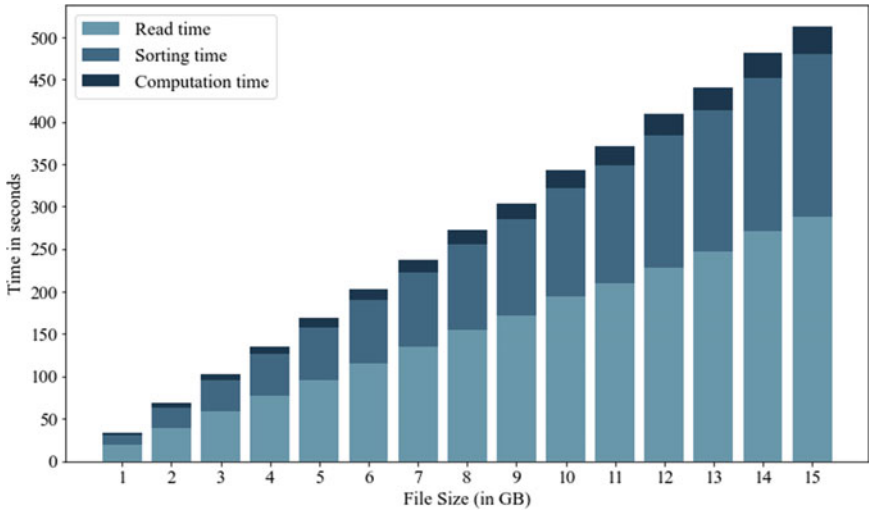


Fig. 4 Time taken by different parts of the code in Hybrid CP algorithm

across data nodes results in a constant time of data transfer, thus resulting in a shorter overall computation time. This functionality is not supported in a GPU architecture as it is.

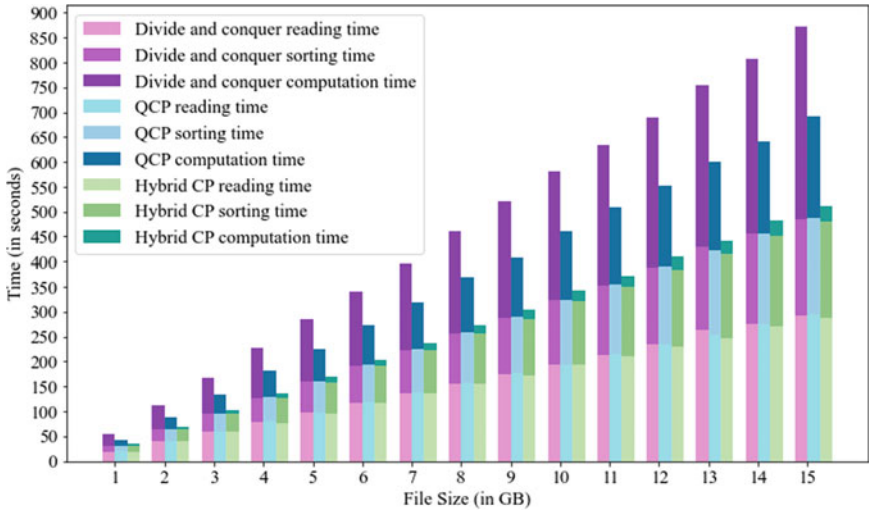


Fig. 5 Time taken by different parts of the code in all the algorithm

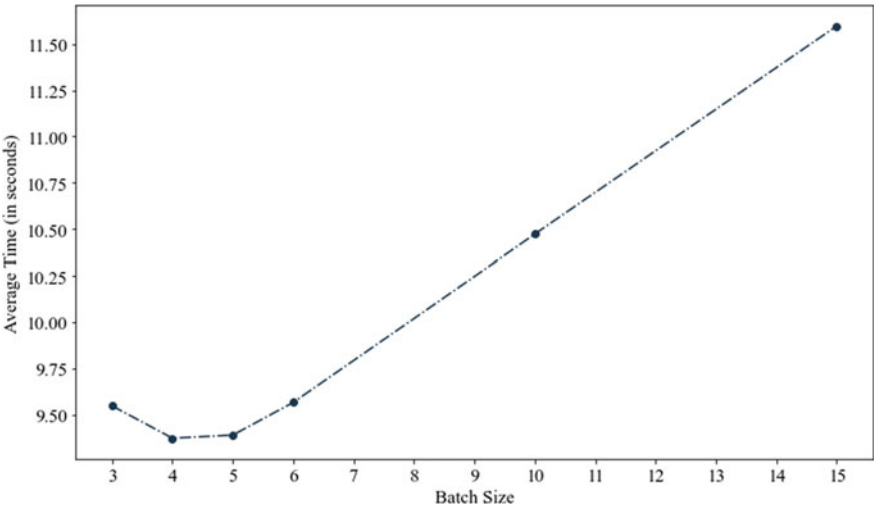


Fig. 6 Average time taken by Hybrid CP algorithm for different batch sizes

5 Conclusion and Future Prospects

Several algorithms have been tested and optimized with GPU and Hybrid configurations as well. The related work section has elaborately covered these methodologies. But closest pair computations are not yet optimized with such techniques. Considering point data sets that mimic the real-world data, HCP is proposed as hybrid

algorithm to compute closest pair. In this work, three CP algorithms are analyzed on CPU and CPU-GPU combinations. With the proposed algorithm on the CPU-GPU hybrid model, up to 12 times speedup in the computation phase is achieved. This translated in up to 91% reduction in the overall execution time. The scope of our experiment does not handle optimization of memory transfers, hence the test cases are limited to 15 GB (beyond which memory corruption had occurred). As future scope higher file sizes are to be tested with alternative memory allocation/access techniques to optimize the data transfer overhead. Suitable algorithms that perform the sorting of the points on the GPU may also help in optimizing the computation.

Acknowledgements We extend our sincere thanks to Prof. N. S. Kumar for giving us insights into our code base that helped us to analyze the overheads of the data transfers and caching.

References

1. Bellini P, Nesi P, Paolucci M, Zaza I (2018) Smart city architecture for data ingestion and analytics: processes and solutions. In: 2018 IEEE fourth international conference on big data computing service and applications (BigDataService). IEEE, pp 137–144
2. Gupta A, Panagiotopoulos P, Bowen F (2020) An orchestration approach to smart city data ecosystems. Technol Forecasting Social Change 153:119929
3. Chan TM, Rahul S, Xue J (2020) Range closest-pair search in higher dimensions. Comput Geometry 91:101669
4. Shan J, Zhang D, Salzberg B (2003) On spatial-range closest-pair query. In: International symposium on spatial and temporal databases. Springer, Heidelberg, pp 252–269
5. Cai X, Zhou S, Rajasekaran S (2018) Jump: a fast deterministic algorithm to find the closest pair of subsequences. In: Proceedings of the 2018 SIAM international conference on data mining. Society for Industrial and Applied Mathematics, pp 73–80
6. Cai X, Mamun AA, Rajasekaran S (2018) Efficient algorithms for finding the closest *l* l-Mers in biological data. IEEE/ACM Trans Comput Biol Bioinformatics 16(6):1912–1921
7. Bespamyatnikh SN (1998) An optimal algorithm for closestpair maintenance. Discrete Comput Geometry 19(2):175–195
8. Mesrikhani A, Farshi M, Iranfar B (2021) A progressive algorithm for the closest pair problem. Int J Comput Math Comput Syst Theor 6(2):130–136
9. Dalkılıç ME, Ergun S (2013) A simple yet fast algorithm for the closest-pair problem using sorted projections on multi-dimensions. In: Information sciences and systems 2013. Springer, Cham, pp 23–34
10. Jiang M, Gillespie J (2007) Engineering the divide-and-conquer closest pair algorithm. J Comput Sci Technol 22(4):532–540
11. Cs.toronto.edu (2022) [online] Available at: http://www.cs.toronto.edu/~vassos/teaching/c73/handouts/closest-pair-of-points.pdf. Accessed 26 Apr 2022
12. Daescu O, Teo KY (2020) Two-dimensional closest pair problem: a closer look. Discrete Appl Math 287:85–96
13. Singh DP, Joshi I, Choudhary J (2018) Survey of GPU based sorting algorithms. Int J Parallel Programming 46(6):1017–1034
14. Manca E, Manconi A, Orro A, Armano G, Milanesi L (2016) CUDA-quicksort: an improved GPU-based implementation of quicksort. Concurrency Comput Pract Experience 28(1):21–43
15. Peters H, Schulz-Hildebrandt O, Luttenberger N (2010) Parallel external sorting for CUDA-enabled GPUs with load balancing and low transfer overhead. In: 2010 IEEE International

Symposium on Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW). IEEE, pp 1–8
16. Eldawy A, Li Y, Mokbel MF, Janardan R (2013) CG_Hadoop: computational geometry in MapReduce. In: Proceedings of the 21st ACM SIGSPATIAL international conference on advances in geographic information systems, pp 294–303
17. Qi M, Yan K, Zheng Y (2019) GPredicates: GPU implementation of robust and adaptive floating-point predicates for computational geometry. IEEE Access 7:60868–60876
18. Doraiswamy H, Freire J (2020) A gpu-friendly geometric data model and algebra for spatial queries. In: Proceedings of the 2020 ACM SIGMOD international conference on management of data, pp 1875–1885
19. Dafir Z, Lamari Y, Slaoui SC (2021) A survey on parallel clustering algorithms for big data. Artif Intelli Rev 54(4):2411–2443
20. Gowanlock M, Rude CM, Blair DM, Li JD, Pankratius V (2018) A hybrid approach for optimizing parallel clustering throughput using the GPU. IEEE Trans Parallel Distributed Syst 30(4):766–777
21. Zhu Z, Xu S, Tang J, Qu M (2019) Graphvite: a high-performance cpu-gpu hybrid system for node embedding. In: The world wide web conference, pp 2494–2504
22. Lee C, Ro WW, Gaudiot JL (2014) Boosting CUDA applications with CPU-GPU hybrid computing. Int J Parallel Programming 42(2):384–404
23. Ryoo S, Rodrigues CI, Baghsorkhi SS, Stone SS, Kirk DB, Hwu WMW (2008) Optimization principles and application performance evaluation of a multithreaded GPU using CUDA. In: Proceedings of the 13th ACM SIGPLAN symposium on principles and practice of parallel programming, pp 73–82
24. Mokhtari R, Stumm M (2014) BigKernel–high performance CPU-GPU communication pipelining for big data-style applications. In: 2014 IEEE 28th international parallel and distributed processing symposium. IEEE, pp 819–828
25. Karim MZ, Akter N (2011) Optimum partition parameter of divide-and-conquer algorithm for solving closest-pair problem. arXiv preprint arXiv:1111.2422
26. Dellosa R, Fajardo A, Medina R (2019) Modified brute force algorithm to solve the closest pair of points problem based on dynamic warping. Indonesian J Electrical Eng Comput Sci 15(3):1629–1636
27. Bright J (2019) Closest pair optimization on modern hardware

Classification of Pap Smear Image of Cervix Cell Using Machine Learning Techniques and Transfer Learning-Based Convolutional Neural Network Architecture and Scrutinizing Their Performances

Mazedul Haque Bhuiyan and Muhammad Ashfakur Rahman Arju

Abstract Different cervical pap smear cell categorization schemes have recently been presented, the majority of which were binary classifications of normal and abnormal cells. This paper presents the findings of a comprehensive investigation on machine learning and deep learning algorithms for binary and multi-class classification on pap smear images from the Herlev dataset. There are 917 photos in this collection, divided into seven normal and pathological categories. The Google Colab platform was used to generate models utilizing all of the techniques using scikit learn and the keras library from TensorFlow. To begin, several repetitions of processes such as feature importance selection, data normalization, standardization, PCA, T-SNE, and others have been imposed on models such as SVM and XGBoost in this work for machine learning approaches. Second, it was demonstrated in this work that a transfer learning-based CNN model from deep learning can outperform machine learning models in terms of binary and multi-class classifications. Furthermore, it was discovered in this work how computationally time efficient it is to apply a transfer learning model, which required roughly 25 min for 100 epochs. Finally, with several iterations of processes and outcomes, this work demonstrates that given enough data for a multi-class pap smear image classification system, the transfer learning CNN model has a higher potential to get the best results than the machine learning models used.

Keywords Pap smear images · CNN · Machine learning · Transfer learning

M. H. Bhuiyan (✉)
BRAC University, 66, Dhaka 1212, Bangladesh
e-mail: mazedul.haque.bhuiyan@g.bracu.ac.bd

M. A. R. Arju
Daffodil International University, Daffodil Smart City, Ashulia, Dhaka, Bangladesh
e-mail: muhammad25-631@diu.edu.bd

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1_75

1 Introduction

Cervical cancer is the world's fourth most prevalent malignancy in women [1]. More than 570,000 women were diagnosed with cervical cancer in 2018, according to the World Health Organization (WHO), and over 300,000 women die from cervical cancer each year [2]. Image analysis and processing techniques have solved numerous difficulties in medical science due to the importance of finding a connected diagnosis and a cure. Basically, image processing combined with machine learning or deep learning has been used to diagnose a variety of malignancies, including breast cancer, lung cancer, and others. The malignant cell can be classified and the corresponding malignancy identified using images or text as a dataset [3]. To have a better automated disease classification result, a good dataset needs to be created first. In our case, for the classification of cervical cancer, we used the Harlev dataset which has been created from pap smear images of cervix cells. The images were initially annotated in order to utilize this dataset in machine learning and deep learning methodologies. The entire dataset is divided into seven classes, which we will go over in detail in our dataset section. We use the data comprising the feature extracted from the pap smear image dataset by a team of graduate thesis students to have an analysis of pap smear image data [4] using Matlab for our machine learning approach. Various machine learning techniques and CAD systems, such as K-nearest neighbors [5, 6], support vector machines (SVM) [6, 7], and artificial neural networks (ANN) [8], have been utilized to detect and classify aberrant cell pictures throughout the last decade. Plissiti and Nikou [9] used the K-PCA (Gaussian kernel) method to classify cervical cell pictures into normal and pathological groups based on nucleus features while ignoring cytoplasmic features. Using seven nucleus features, some writers found the best harmonic mean (H-mean) classification sensitivity of 90.58%. Chankong et al. [8] segmented cervical cell pictures into the cytoplasm, nucleus, and backdrop using a patch-based fuzzy C-means (FCM) clustering technique. To categorize cervical cell pictures, the segmented images were put into a variety of classifiers, K-nearest neighbor, including Bayesian, linear discriminant analysis, SVM, and artificial neural networks. Using the ERUDIT dataset to categorize cell pictures into two classifications, they achieved the maximum accuracy of 97.83%. For two-class issues, the authors employed the Herlev and LCH datasets and achieved the greatest accuracy of 99.27%. For a private dataset with 14 textures and 30 form features, Ashok and Aruna [7] employed a multi-thresholding approach for segmentation and an SVM classifier with 99% accuracy. A different dataset with 150 photos was used to obtain the characteristics. Despite the fact that the accuracy level is higher, no information about the intricate and larger dataset is available. For an exhaustive comparison of five classifiers and FS, Chankong et al. [10] employed artificial neural networks (ANNs). A three-layer backpropagation ANN with nine features produced the greatest results. It was divided into 2, 4, and 7 classes. 7 classes have a 93.78% accuracy and a 96.69% specificity. Zhang et al. [11] developed a deep convolutional neural network for pap smear picture classification with a size of (256×256 pixels). The suggested method employs transfer learning to efficiently set the weights of deep

neural networks using the ImageNet dataset. To our understanding, this model has a 98% accuracy rate, with the highest accuracy for pap smear classification research. Unlike machine learning, feature extraction in deep learning is an automated process. This property shortens the time it takes to create a model. However, the typical training time is slow, taking 4 h. In this investigation, each training effort consisted of classifying a pap smear image of cervix cell 3 for 30 epochs with a 256 $\times$ 256 pixels image size. In a nutshell, this research compares the performance of machine learning and deep learning algorithms for binary and multi-class categorization of cervical cell pictures. The importance of this work is to demonstrate how models are tested and improved through optimization and a comparison of the accuracy of various methodologies. Deep learning techniques can be used to find insights in complicated big data, starting with data preparation, training the model, and evaluating it [12]. We started with the feature extracted data, but later, we implement the PCA for normalization and t-SNE for standardization. Sequentially, the work is iterated multiple times with different approaches to see the accuracy coming out from 3 different classes. We divided the classes by 2, 5, 7 during the iteration, with the classes consisting of both cancerous abnormal and noncancerous cervix cell images. We also assessed XGBoost, an SVM classifier, using performance metrics taken from the feature extraction dataset. Moreover, this work was done on Google platform, and it took relatively less computational timing for training machine learning models and other feature engineering for those. Though in this study a transfer learning model is used, it took 25 min for 100 epochs of training for the deep learning model where the image size was 224 $\times$ 224 pixels which is pretty amazing considering the computational time of others using any costly hardware.

2 Dataset Description and Extracted Feature Preprocessing

Our first goal was to use the classifiers on a main dataset obtained in our own region, but after putting in a lot of time and effort, we discovered that there were no adequate picture datasets of cervix cells to work with within any hospitals or institutes that we could access. So, we used the well-known benchmark dataset on this topic from Herlev University Hospital in Denmark, which is all freely available on the Internet. They used a digital camera and a microscope to acquire a dataset of single cells from the cervix area. Regardless of the hospital's practical and financial limits, the database contains the most accurate diagnosis possible. As a result, several researchers have used it as a reference database. Each cell was manually sorted into one of seven groups and examined by two cytotechnicians; complex samples were checked by doctors and rejected if there were any disagreements. The dataset consists 917 samples that are spread unevenly over seven classes, with noncancerous cases accounting for 242 and cancerous instances accounting for 675. The noncancerous types are classified as Class One, Two, and Three, and these samples come from a variety of cell types, including superficial and intermediate squamous epithelial, as well as columnar epithelial, with 74, 70, and 98 cases, respectively. Class Four, Five, Six,

and Seven are carcinogenic forms that include mild, moderate, and severe squamous non-keratinizing dysplasia, as well as squamous cell carcinoma in situ intermediate cell types, with 182, 146, 197, and 150 cases, respectively. Classifiers are frequently employed to distinguish normal from pathological cells or vice versa, resulting in a two-class problem. Whereas in our case, we tried to see how the classifiers perform in a multi-class problem. We needed the extracted features from the cell photos for the machine learning approach, which we acquired along with the database. The authorities from Herlev University segmented the photos using a commercial software tool called CHAMP (Dimac). The retrieved features are made up from the combination of features from segmented cell pictures and non-segmented cell pictures. Twenty attributes are retrieved from the photographs of each cell to describe every sample. Shortest and longest diameter, elongation and roundness of nucleus and cytoplasm, perimeter of nucleus and cytoplasm, nucleus position and maxima, minima of nucleus and cytoplasm are among the attributes. The features were then submitted to feature engineering to determine if we might increase our performance, which we will discuss in the next part.

3 Methodology

In this section, how we approached our result, the algorithms, the feature engineering methods, we have followed during our work, and the important hyperparameters selection criteria's are explained in detail. There are two different machine learning algorithm such as SVM, XGBoost, and a transfer learning-based deep neural network (CNN); we have used for classification, and they are presented here. Each algorithm and dataset has different hyperparameters and effects in the model. We have used the most used dataset on this domain, which is the Herlev dataset to compare the methodologies and hyperparameters applied here. This dataset contains 20 structural characteristics with labels that are relevant for the machine learning algorithms utilized. Since this data are multidimensional, we have applied some feature engineering methods like XGBoost for feature selection, Pca, t-SNE, standardization, normalization, and Py-mrmr here so that we can see if it can make any changes to the results using the outcomes from these methods. However, we have applied each algorithm for three-tier multi-class problem for instance at first we have tried the problems for two-class problem keeping the first three1, 2, 3 normal noncancerous classes as normal class and the last four 4, 5, 6, 7 abnormal cancerous classes as abnormal class; then, we have tried the classification algorithms for a multiple class problem of five classes; here, we kept the first three normal noncancerous classes as a single normal class and last four abnormal cancerous classes as four individual single classes each, and finally, we have tried the algorithms for seven classes classification problem keeping the seven different normal and abnormal classes separate as they are. Besides this three-tier multiple class classification, we have developed our workflow in five iterative steps which is shown in Fig. 1. To have a short brief about our steps we can say, at first in our first iteration, we applied the XGBoost and

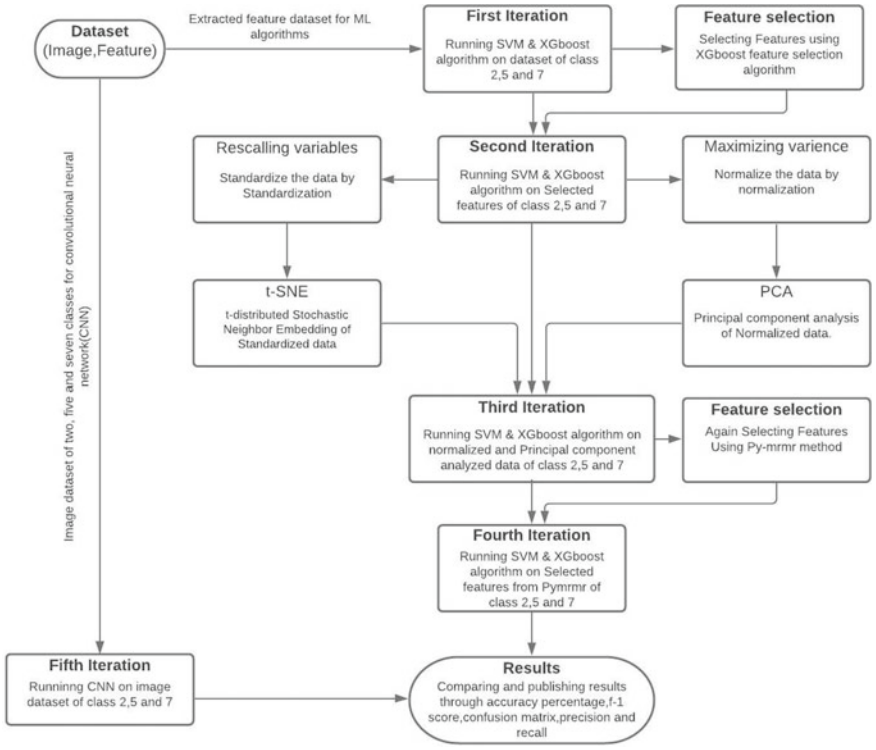


Fig. 1 Workflow of how machine learning and deep learning algorithms are used to diagnose cervical cancer from multi-class pap smear images

SVM algorithm on the dataset of two, five, and seven classes. In our second iteration, we have applied the XGBoost feature selection algorithm on the featured dataset and selected the features based on a threshold value, and then, we again followed the iteration one on the selected features. However, in our third iteration, we have applied normalization, standardization, Pca, and t-SNE, respectively, on our dataset as our feature engineering purpose; then, we have applied the machine learning algorithms again on the acquired data of two, five, and seven classes. In our fourth iteration, we have applied Py-mrMR for feature selection again and then applied the algorithms for two, five, and seven classes like before to see the changes in results using the selected features by Py-mrMR. Finally, in our fifth iteration, we have applied our transfer learning-based convolutional neural network algorithm (CNN) on the image dataset, like before we have also separated the image dataset into two, five, and seven classes and used that for classification.

3.1 Feature Engineering

Multiple feature engineering processes are used to optimize the dataset of features extracted from the images so that it gives a better enhanced result in machine learning algorithms. Feature selection is a part of feature engineering and an important component of machine learning since the feature must be unique in order to increase the likelihood of better results. XGBoost and Py-mrmr are such feature-selecting algorithms that have been used in this case. To achieve superior results, all differences between measurements for any given cell class must be quasi and detectable, which necessitates tweaking the data variances on occasion. To pursue such ambition techniques like normalization, PCA, standardization, and t-SNE have been used.

3.1.1 XGBoost-Based Feature Selection

Feature selection has long been a challenging but significant problem in classification. This type of situation necessitates feature selection techniques that can aid classifiers in not just improving classification accuracy, but also reducing redundant features to the greatest extent possible. As a result, a new wrapped feature selection technique, XGBSFS, was presented to better address feature selection in classification issues. To avoid the limitations of a single importance metric, the thought process of generating trees in XGBoost was utilized as a reference, and the importance of features from three importance measures was quantified. The feature subset was then searched using improved sequential floating forward selection (ISFFS) to ensure that it was of good quality [13].

3.1.2 Py-mrMR Feature Selection Approach

The mRMR is used as a technique for selecting features which prioritizes features with a high association with the class (output) and a weak association with each other. For continuous qualities, the test statistic for F-test can be used to quantify the connection with the class (relative importance), whereas for discrete features, the Pearson correlation coefficient can be used to analyze the relationships between features (redundancy). Then, using a greedy search to optimize the objective function, which is a mechanism of relevancy and redundancy, each characteristic is chosen one at a time. The MIQ ("Mutual Information Quotient criterion") and the MID ("Mutual Information Difference criterion") are two types of algorithms that indicate the variance or ratio of relevancy and redundant, respectively [14].

3.1.3 Normalization

Normalization technique is used in data processing to turn the data into a set of values. The efficiency and accuracy of classifiers may be harmed as a result of this transition. Although neural network classifiers may gain considerably by normalizing the data, others, like decision tree classifier and Naive Bayesian, may not. Because distance metrics can produce meaningful values after normalization, algorithms that use distance measures may produce appropriate results with normalization. Furthermore, because normalization speeds up the training step, it has the potential to improve the performance of the classifier of neural networks (NNW) [15]. For our dataset, we investigate the effects of z-score normalization as well as min–max normalization on classifiers.

3.1.4 Principal Component Analysis (PCA)

Throughout many categorization circumstances, several of the dataset's attributes lack distinguishing traits that allow them to be separated. Some characteristics may be highly connected, while others may be utterly unrelated to the sample. As a result, data reduction strategies are available to remove these ineffective elements. PCA is a well-known method for reducing dimensionality [16]. Using orthogonal transformation, it converts the entire dataset into a new facility of lineally uncorrelated properties.

3.1.5 Standardization

When features in an input dataset have considerable discrepancies between their ranges or simply when they are measured in multiple measurement units, standardization is required. Many machine learning algorithms struggle because of these discrepancies in initial feature ranges. For example, in models that use distance computation, if one of the features has a wide range of vales, the distance will be determined by that feature. Because we have features with data that are reasonably near in range and indifferent in measurement unit standardization, machine learning models were not expected to perform well. So, after comparing the normalization and standardization results, we decided to continue forward with normalization and PCA rather than standardization and t-SNE, and for standardization, we used the Z-score approach in our situation.

3.1.6 t-Distributed Stochastic Neighbor Embedding

SNE generates decent visuals, but it suffers from a tough-to-evaluate cost objective and a issue we call the "crowding problem." To overcome these challenges, a novel technique known as "t-Distributed Stochastic Neighbor Embedding," or "t-SNE,"

has been created. There are two ways that differ the usage of cost objective between t-SNE and SNE: (1) A flatten kind of the SNE cost objective with uncomplicated gradients is used by it which was concisely introduced by Cook et al., and (2) rather than using Gaussian distribution in low-dimensional space to compute similarity between two points, it uses student-t-distribution. t-SNE uses an intense dispersion in low-dimensional space to handle both the crowding issue and the optimization issues of SNE [17].

3.2 Machine Learning

In many fields, such as regression and classification, traditional machine learning techniques are applied. To create a model, machine learning algorithms use data attributes. The model's performance for classification with the highest metrics improves as a result of the distinguishing features. The nucleus and cytoplasm areas, as well as their ratios, brightness, shortest and longest diameters, elongation, perimeter, location, maxima, minima, and roundness, are all significant qualities of cervical cells. The features of the dataset of two, five, and seven classes are utilized to classify normal and abnormal cervical cells using SVM with varying kernels and XGBoost classifiers.

3.2.1 Support Vector Machine Classifier

SVMs are statistical learning theory-based supervised machine learning algorithms. SVMs are utilized in numerous fields such as the face, optical character, and speech recognition, and they have several advantages such as high efficiency for a large dataset [18]. SVM's kernel is provided as a radial basis function (rbf) for the maximum accuracy because of the advantages of kernels.

3.2.2 XGBoost Classifier

XGBoost is a popular and relatively new machine learning algorithm that has proven to be quite efficient, winning 50% of the challenges on machine learning platforms such as Kaggle [19]. XGBoost is a gradient boosted technique that is based on decision trees. Tree pruning, parallelization, and cross-validation are just a few of the systems and algorithmic optimizations it has [20].

3.3 Deep Learning

The robust feature extraction and information mining capabilities of the convolutional neural network (CNN) have gotten a lot of attention. Because of its powerful feature extraction and learning process, CNN has been employed for a range of applications such as object recognition, picture super-resolution, semantic segmentation, and so on. Various CNN designs were presented to increase the respective system performance by keeping the baseline learning topology unchanged. AlexNet, VGG16, and VGG19 are three well-known CNN architectures for object recognition applications [21]. For our multi-class picture classification, we fine-tuned and employed the state-of-the-art CNN architecture VGG16. After experimenting with a variety of CNN models, we discovered that transfer learning as VGG16 produces the best results. We didn't retrain the main model because of our tiny dataset and significant data similarities; instead, we altered the output layer to match our classes and fine-tuned the optimizers to improve performance. In addition, we used the pre-trained model to extract features from pap smear images. Table 1 lists the properties of the neural network and dataset used in the model. In addition, an overview of the VGG16 CNN model architecture is provided (see Fig. 2).

Table 1 CNN model parameters

Property	Value
Image size	(224,224,3)
2D pooling size	2 × 2
Convolutional layer number	13
Epoch number	100
Test data/total data	0.05
Validation data/training data	0.13

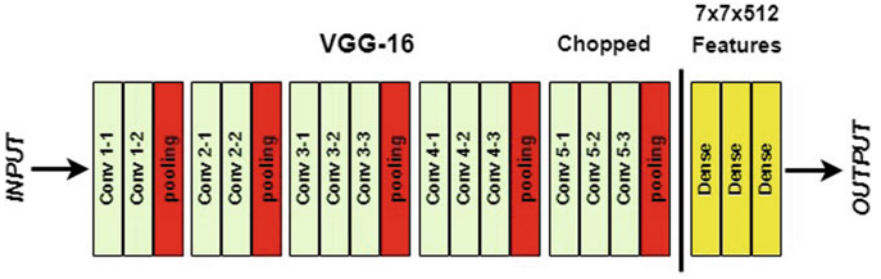


Fig. 2 This diagram depicts the VGG16 convolutional neural network model's fundamental architecture for pap smear image classification

3.4 Metrics

Metrics are used to compare and measure the performance of different models. The metrics, which are generated using a recall, specificity, confusion matrix, precision, reveal accuracy, and F_1 score. Confusion matrix has the essential parameter of metrics included in it. The metrics can be calculated using necessary parameters, and those are A true-positive (TP) result, the true-negative (TN) number, the false-positive (FP) value, and the false-negative (FN) value. The performance metrics provide additional information regarding models produced by the TP, TN, FP, and FN values. The rate of genuine samples among all samples is the accuracy. A higher rate of correctly identified data indicates a higher level of accuracy. Moreover, the rate of accurately predicted positive samples among all positive predicted samples is the precision value. For actual results, better precision equals more correctly identified data. Furthermore, the recall value is the percentage of accurately predicted positive samples among all positive samples that are actually positive. Whereas, the F_1 score is the sum of recall and precision. It stops you from selecting the incorrect model unless the dataset is split correctly.

- Accuracy = True Positive + True Negative/True Positive + False Positive + True Negative + False Negative
- Precision = True Positive/True Positive + False Positive
- Recall = True Positive/True Positive + False Negative
- F_1 Score = 2 * (Precision * Recall)/(Precision + Recall)

4 Result

The model's efficacy for cervical cancer classification from pap smear images is obtained in this study. A table is used to compare different models quickly. This section also highlights which models outperformed the others in terms of performance. In Table 2, all of the results for several classes of the dataset of model metrics are proposed and compared. Since it was the primary concern of this work to find out that which model performs better for multi-class data of ours regardless of having a very small dataset, here, we will see how the model performed for class 2, 5, and 7. Firstly, in machine learning technique, SVM has 94% model accuracy with 93% F_1 score 91% precision and 90% recall value for class 2 after first iteration which makes it the best performed model in this section even though the XGBoost model has also gained 94% accuracy after second iteration of our process but considering the F_1 score 92%, precision 92% and recall value of 89%. Secondly, for class 5 and class 7, the numbers have dropped down in all metrics which was predictable due to insufficient data, but even there we wanted to see how the both machine learning model performs and here too the SVM outperformed the decision tree model. After the first iteration, the SVM model had 64% and 63% accuracy for class 5 and 7, respectively,

Table 2 Machine and deep learning approaches' evaluation metrics

		First iteration		Second iteration		Third iteration		Fourth iteration		Fifth iteration
		SVM	Xgboost	SVM	Xgboost	SVM	Xgboost	SVM	Xgboost	CNN
Accuracy (%)	2 Class	94	93	94	91	94	86	93	90	94
	5 Class	64	67	62	62	64	49	57	59	66
	7 Class	63	64	61	59	64	48	56	59	64
F_1-score (%)	2 Class	93	93	91	87	91	80	89	85	94
	5 Class	63	66	59	59	61	46	54	56	60
	7 Class	63	63	63	62	67	53	60	61	55
Precision (%)	2 Class	91	92	91	86	91	79	90	85	75
	5 Class	61	62	59	59	61	46	54	56	67
	7 Class	67	65	63	63	67	51	59	61	61
Recall (%)	2 Class	90	89	90	88	90	82	89	86	95
	5 Class	62	62	60	60	62	47	55	57	58
	7 Class	68	63	65	61	68	54	60	62	56

with the same F_1 score of 63% where the precision value was 61% and 67% with the recall value of 62% and 68%, respectively, whereas even the XGBoost model had a better accuracy 67% for class 5 and 64% for class 7, but again considering the overall recall value for both classes, it falls a little behind than the SVM model. Moreover, going through all the iteration it was visible that the Decision tree model was not consistent in its performance metrics Whereas the Support vector machine seemed to maintain a consistent result in all classes which makes it the best model here in machine learning session. Now, coming to deep learning, the transfer learning CNN model performed far better than any of the machine learning models applied here. It is also important to mention that this result here for this model is from the test data, so it is so clear how well the fine-tuned CNN model performed here for class 2 with 94% accuracy and F_1 score 75% precision and 94% recall. Despite having the minimal dataset, the CNN model has gained 66% accuracy, 60% F_1 score,67% precision, and 58% recall. Moreover, for class 7, the CNN model has 64% accuracy, 55% F_1 score, 61% precision, and 56% recall. Despite having a not so good result for class 5 and class 7 with the consistent performance in all metrics, it is expected that with a sufficient amount of data, this model will perform much better than the machine learning models for multi-class scenarios.

5 Conclusion

Machine learning and deep learning approaches are employed, analyzed, and compared in this study to see which method performs best for the Harlev dataset, not only for binary but also for multi-class classification. It is true that by fine-tuning hyperparameters, a model or algorithm can operate better for a dataset and produce a great result. However, in order to get a better result from a model, it is also necessary to have a well-balanced dataset. Both the hyperparameters optimization and the dataset are essential for obtaining the best results from any algorithm or model. Furthermore, it has been demonstrated that CNN can extract better features from data than handcrafted features, resulting in an astounding result. We also believe that, with a balanced dataset for each significant class and better hyperparameter optimization, the CNN model will outperform machine learning techniques for various classes, as demonstrated in this paper.

References

1. Khazaei Z, Jarrahi AM, Momenabadi V, Ghorat F, Adineh H, Sohrabivafa M, Goodarzi E (2019) Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide stomach cancers and their relationship with the human development index (HDI). World Cancer Res J 6(9)
2. Okunade KS (2020) Human papillomavirus and cervical cancer. J Obstet Gynaecol 40(5):602–608

3. Ayaka Iwai TT (2017) Automatic diagnosis supporting system for cervical cancer using image processing. In: 2017 56th Annual conference of the society of instrument and control engineers of Japan (SICE). IEEE, pp 479–482
4. Jantzen J, Dounias G (2006) Analysis of pap-smear image data. In: Proceedings of the nature-inspired smart information systems 2nd annual symposium, vol 10, pp 1–11
5. Sharma M, Singh SK, Agrawal P, Madaan V (2016) Classification of clinical dataset of cervical cancer using KNN. Indian J Sci Technol 9(28):1–5
6. Kumar R, Srivastava R, Srivastava S (2015) Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. J Med Eng 2015
7. Ashok B, Aruna P (2016) Comparison of feature selection methods for diagnosis of cervical cancer using SVM classifier. Int J Eng Res Appl 6:94–99
8. Chankong T, Theera-Umpon N, Auephanwiriyakul S (2009) Cervical cell classification using Fourier transform. In: 13th International conference on biomedical engineering. Springer, Berlin, pp 476–480
9. Plissiti ME, Nikou C (2012) Cervical cell classification based exclusively on nucleus features. In: International conference image analysis and recognition. Springer, Berlin, pp 483–490
10. Chankong T, Theera-Umpon N, Auephanwiriyakul S (2014) Automatic cervical cell segmentation and classification in pap smears. Comput Methods Programs Biomed 113(2):539–556
11. Zhang L, Lu L, Nogues I, Summers RM, Liu S, Yao J (2017) Deeppap: deep convolutional networks for cervical cell classification. IEEE J Biomed Health Inf 21(6):1633–1643
12. Jan B, Farman H, Khan M, Imran M, Islam IU, Ahmad A, Ali S, Jeon G (2019) Deep learning in big data analytics: a comparative study. Comput Electr Eng 75:275–287
13. Zhanshan L, Zhaogeng L (2019) Feature selection algorithm based on xgboost. J Commun 40(10):101
14. Radovic M, Ghalwash M, Filipovic N, Obradovic Z (2017) Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. BMC Bioinf 18(1):1–14
15. Al Shalabi L, Shaaban Z, Kasasbeh B (2006) Data mining: a preprocessing engine. J Comput Sci 2(9):735–739
16. Tan PN, Steinbach M, Kumar V (2006) Data mining introduction. People's Posts and Telecommunications Publishing House, Beijing
17. Van der Maaten L, Hinton G (2008) Visualizing data using t-sne. J Mach Learn Res 9(11)
18. Cortes C, Vapnik V (1995) Support vector machine. Mach Learn 20(3):273–297
19. Nielsen D (2016) Tree boosting with xgboost-why does xgboost win "every" machine learning competition? Master's thesis, NTNU
20. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, Chen K et al (2015) Xgboost: extreme gradient boosting. R Package Version 0.4-2 1(4):1–4
21. Shaha M, Pawar M (2018) Transfer learning for image classification. In: 2018 Second international conference on electronics, communication and aerospace technology (ICECA). IEEE, pp 656–660

Author Index

S. Fong et al. (eds.), *ICT Analysis and Applications*, Lecture Notes in Networks and Systems 517, https://doi.org/10.1007/978-981-19-5224-1