Lecture Notes in Networks and Systems 399

Xiao-Zhi Gao · Shailesh Tiwari ·
Munesh C. Trivedi ·
Pradeep Kumar Singh ·
Krishn K. Mishra *Editors*

Advances in Computational Intelligence and Communication Technology

Proceedings of CICT 2021

Springer

Lecture Notes in Networks and Systems

Volume 399

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).

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

Xiao-Zhi Gao · Shailesh Tiwari ·
Munesh C. Trivedi · Pradeep Kumar Singh ·
Krishn K. Mishra
Editors

Advances in Computational Intelligence and Communication Technology

Proceedings of CICT 2021

Springer

Editors
Xiao-Zhi Gao
University of Eastern Finland
Kuopio, Finland

Munesh C. Trivedi
National Institute of Technology Agartala
Agartala, Tripura, India

Krishn K. Mishra
Motilal Nehru National Institute
of Technology
Allahabad, Uttar Pradesh, India

Shailesh Tiwari
Krishna Engineering College
Ghaziabad, Uttar Pradesh, India

Pradeep Kumar Singh
Krishna Institute of Engineering
and Technology
Ghaziabad, Uttar Pradesh, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-16-9755-5 ISBN 978-981-16-9756-2 (eBook)
https://doi.org/10.1007/978-981-16-9756-2

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

Preface

The International Conference CICT 2021 is a major multidisciplinary conference organized to provide a forum for researchers, educators, engineers, and government officials involved in the general areas of computational intelligence and communication technology to disseminate their latest research results, exchange views on the future research directions of these fields, exchange computer science, integrate its practice and application of the academic ideas, improve the academic depth of computer science and its application, and provide an international communication platform for educational technology and scientific research for the universities and engineering field experts and professionals.

Nowadays, globalization of academic and applied research is growing with speedy pace. Computer, communication, and computational sciences are the heating areas with lot of thrust. Keeping this ideology in preference, Janardan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India, has come up with an event—International Conference on Computational intelligence and Communication Technology (CICT 2021) during October 29–30, 2021.

This is the sixth time the International Conference on Computational Intelligence and Communication Technology (CICT 2021) is organized, with a foreseen objective of enhancing the research activities at a large scale. The Technical Program Committee and Advisory Board of CICT 2021 include eminent academicians, researchers, and practitioners from abroad as well as from all over the nation.

A sincere effort has been made to make it an immense source of knowledge for all and includes 41 manuscripts. The selected manuscripts have gone through a rigorous review process and are revised by authors after incorporating the suggestions of the reviewers.

CICT 2021 received around 290 submissions from around 842 authors of 6 different countries such as American Samoa, Estonia, Greece, Iran, Taiwan, and Vietnam. Each submission has been gone through the plagiarism check. On the basis of plagiarism report, each submission was rigorously reviewed by at least two reviewers with an average of 2.9 per reviewer. Even some submissions have more than two reviews. On the basis of these reviews, 41 high-quality papers were selected for publication in this proceedings volume, with an acceptance rate of 14.1%.

We are thankful to the keynote speakers—Prof. Andrew K. Ng, SIT, Singapore, and Dr. K. K. Mishra, MNNIT Allahabad, India, to enlighten the participants with their knowledge and insights. We are also thankful to delegates and the authors for their participation and their interest in CICT 2021 as a platform to share their ideas and innovation. We are also thankful to Prof. Dr. Janusz Kacprzyk, Series Editor, LNNS, Springer, and Mr. Aninda Bose, Senior Editor, Springer Nature, for providing guidance and support. Also, we extend our heartfelt gratitude to the reviewers and Technical Program Committee Members for showing their concern and efforts in the review process. We are indeed thankful to everyone directly or indirectly associated with the conference organizing team leading it toward the success.

Although utmost care has been taken in compilation and editing, however, a few errors may still occur. We request the participants to bear with such errors and lapses (if any). We wish you all the best.

Udaipur, Rajasthan, India

Xiao-Zhi Gao
Shailesh Tiwari
Munesh C. Trivedi
Pradeep Kumar Singh
Krishn K. Mishra

Contents

Editors and Contributors

About the Editors

Prof. Xiao-Zhi Gao received his B.Sc. and M.Sc. degrees from the Harbin Institute of Technology, China, in 1993 and 1996, respectively. He obtained his D.Sc. (Tech.) degree from the Helsinki University of Technology (now Aalto University), Finland, in 1999. He has been working as Professor at the University of Eastern Finland, Finland, since 2018. He is also Guest Professor at the Harbin Institute of Technology, Beijing Normal University, and Shanghai Maritime University, China. His current research interests are nature-inspired computing methods with their applications in optimization, data mining, machine learning, control, signal processing, and industrial electronics.

Prof. Shailesh Tiwari currently works as Director, Krishna Engineering College, Ghaziabad, India. He is an alumnus of Motilal Nehru National Institute of Technology Allahabad, India. His primary areas of research are software testing, implementation of optimization algorithms and machine learning techniques in various problems. He has published more than 80 publications in International Journals and in Proceedings of International Conferences of repute. He is edited Scopus, SCI, and E-SCI-indexed journals. He has also edited several books published by Springer. He has organized several international conferences under the banner of IEEE and Springer. He is Senior Member of IEEE, Member of IEEE Computer Society, Fellow of Institution of Engineers (FIE).

Dr. Munesh C. Trivedi currently works as Associate Professor in the Department of Computer Science and Engineering, NIT Agartala, Tripura, India. He has published 20 text books and 80 research publications in different international journals and proceedings of international conferences of repute. He has received Young Scientist and numerous awards from different national as well international forums. He has organized several international conferences technically sponsored by IEEE, ACM, and Springer. He is on the review panel of IEEE Computer Society, *International*

Journal of Network Security, *Pattern Recognition Letters* and *Computer & Education* (Elsevier's Journal). He is Executive Committee Member of IEEE UP Section, IEEE India Council, and also IEEE Asia Pacific Region 10.

Prof. Pradeep Kumar Singh is currently working as Professor and Head in the department of CS at Krishna Institute of Engineering and Technology, Ghaziabad, Uttar Pradesh. He is a life membership of CSI, IEI and promoted to a senior member grade from CSI and ACM. He is Associate Editor of the (*IJISMD*), *IJAEC*, IGI Global USA, *SPY*, Wiley *IJISC*, a scientific peer-reviewed journal. He is recently appointed as Section Editor, *Discover IoT*, Springer Journal. He has published 106 research papers in various international journals and conferences. He has edited a total 16 books from Springer and Elsevier. He has Google Scholar citations 944, H-index 18, and i-10 Index 31. His recently published book titled *Handbook of Wireless Sensor Networks: Issues and Challenges in Current Scenario's* from Springer has reached to 11,000 downloads in last few months.

Dr. Krishn K. Mishra is currently working Assistant Professor, Department of Computer Science and Engineering, Motilal Nehru National Institute of Technology Allahabad, India. His primary area of research includes evolutionary algorithms, optimization techniques and design, and analysis of algorithms. He has also published more than 50 publications in international journals and in proceedings of internal conferences of repute. He is serving as a program committee member of several conferences and also edited Scopus and SCI-indexed journals. He is also member of reviewer board of *Applied Intelligence Journal*, Springer.

Contributors

Ankush Agarwal Department of Computer Engineering and Applications, GLA University, Mathura, India

Gaurav Agrawal Inderprastha Engineering College, Ghaziabad, India

Rohit Agrawal Department of Computer Engineering and Application, GLA University, Mathura, India

Ruchi Agrawal Department of Computer Engineering and Applications, GLA University, Mathura, India

A. Ananth Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India

L. Arockiam St. Joseph's College (Autonomous), Tiruchirappalli, India

P. Aruna Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India

P. Baskaran Easwari Engineering College, Chennai, India

Almas Begum Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Brijesh Bharadwaj Dr. RL Avadh University, Ayodhya, India

Diwakar Bhardwaj Department of Computer Engineering and Application, GLA University, Mathura, India

Shweta Chauhan Lloyd Institute of Engineering and Technology, Gr. Noida, U.P, India

Diksha Dani Inderprastha Engineering College, Ghaziabad, India

Devansh Devagya ABES Engineering College, Ghaziabad, India

Aruna Dhamija Institute of Business Management, GLA University, Mathura, India

Shivam Dhar Maharaja Surajmal Institute of Technology, New Delhi, India

A. Cecil Donald Assistant Professor, Department of Computer Science, CHRIST (Deemed to be University), Central Campus, Bangalore, India

Anil Kumar Dubey ABES Engineering College Ghaziabad, Ghaziabad, Uttar Pradesh, India

Rinky Dwivedi Maharaja Surajmal Institute of Technology, New Delhi, India

Neetu Faujdar Department of Computer Engineering and Applications, GLA University, Mathura, India

Vimal Gaur Maharaja Surajmal Institute of Technology, Delhi, India

Harsh Kumar Goel ABES Engineering College, Ghaziabad, India

Anil Kumar Goswami DRDO, Delhi, India

M. Gunasekaran Research Department of Computer Science, Government Arts College, Dharmapuri, India

Aakansha Gupta Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

Asmita Sen Gupta Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

Avadhesh Kumar Gupta USCI, Karnavati University, Gujrat, India

Vandana Kumari Lloyd Institute of Engineering and Technology, Gr. Noida, U.P, India

Nguyen Thi Hang Thai Nguyen University, University of Information and Communication Technology, Hanoi, Vietnam

D. Hemalatha Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Dinh Thi Hien Thai Nguyen University of Sciences, Thai Nguyen, Vietnam

Nguyen Thanh Hoang University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh, Vietnam

Dinh Tran Ngoc Huy Banking University HCMC, Ho Chi Minh city, Vietnam; International University of Japan, Niigata, Japan

Himani Jain ABES Engineering College, Ghaziabad, Uttar Pradesh, India

Naramalli Jaya Krishna School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Nisha Joseph Department of Computer Science and Engineering, SAINTGITS College of Engineering (Autonomous), Kerala, India

Ranjana Joshi Department of CEA, GLA University, Mathura, India

Amrita Jyoti Ajay Kumar Garg Engineering College, Ghaziabad, U.P, India

Ravi Kalra Lloyd Institute of Engineering and Technology, Gr. Noida, U.P., India

Nidhi Kanwar CEA Department, GLA University, Mathura, UP, India

Rahul Katarya Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

R. Kavitha Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India

Irfan Khan Krishna Engineering College, Mohan Nagar, Ghaziabad, Uttar Pradesh, India

Rakesh Kumar Department of CEA, GLA University, Mathura, India

Monika Kumari Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

Toan Linh Vu Le Faculty of Banking and Finance, Van Lang University, Ho Chi Minh City, Vietnam;
School of Banking, University of Economics Ho Chi Minh City (UEH), Ho Chi Minh City, Vietnam

Trung-Hieu Le Dai Nam University, Hanoi, Vietnam

Sachin Malhotra Chandigarh University, Chandigarh, India

Anju Mishra Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

Rashmi Mishra Krishna Engineering College, Ghaziabad, Uttar Pradesh, India; Lloyd Institute of Engineering and Technology, Gr. Noida, U.P, India

Vivek Mishra ABES Engineering College, Ghaziabad, Uttar Pradesh, India

P. Mohan Babu Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India

Divya Mohan Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Narendra Mohan Department of Computer Engineering and Application, GLA University, Mathura, India

S. Naveen Kumar Easwari Engineering College, Chennai, India

Parashu Ram Pal SAGE University, Bhopal, Madhya Pradesh, India

Shruti Pancholi Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India

Sanjeet Pandey MUIT, Lucknow, India

Deepika Pandoi Institute of Business Management, GLA University, Mathura, India

Pankhuri Maharaja Surajmal Institute of Technology, New Delhi, India

Pankaj Pathak Symbiosis Institute of Digital and Telecom Management Symbiosis International (Deemed University), Pune, India

Kavita K. Patil ACS College of Engineering, Bangalore, Karnataka, India

Sunita S. Patil ACS College of Engineering, VTU, Bangalore, India

Duy Khanh Pham Faculty of Banking and Finance, Van Lang University, Ho Chi Minh City, Vietnam;
School of Banking, University of Economics Ho Chi Minh City (UEH), Ho Chi Minh City, Vietnam

A. Y. Prasad SJB Institute of Technology, Bangalore, Karnataka, India

Renu Prasad Department of Computer Engineering and Application, GLA University, Mathura, India

Narayanan Prasanth School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Akash Punhani ABES Engineering College, Ghaziabad, India

Harsh Raj Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

K. Rajathi Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai, Tamil Nadu, India

S. V. Rajiga Research Department of Computer Science, Government Arts College (Autonomous), Salem, India

Rupa Rani Krishna Engineering College, Mohan Nagar, Ghaziabad, Uttar Pradesh, India;
Lloyd Institute of Engineering and Technology, Gr. Noida, U.P., India

R. Rengarajan Easwari Engineering College, Chennai, India

Rishabh Galgotias College of Engineering and Technology, Greater Noida, India

M. Sanjana Sri Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India

Mala Saraswat ABES Engineering College Ghaziabad, Ghaziabad, Uttar Pradesh, India

Harsh Saxena ABES Engineering College, Ghaziabad, India

T. Senthil Kumaran ACS College of Engineering, VTU, Bangalore, Karnataka, India

B. J. Hubert Shanthan Kristu Jayanti College (Autonomous), Bengaluru, India

Ashish Sharma Department of Computer Engineering and Application, GLA University, Mathura, India

Himanshu Sharma Department of Computer Engineering and Application, GLA University, Mathura, India

Shashi Shekhar Department of Computer Engineering and Application, GLA University, Mathura, India

Shivanshi Shokeen Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

Rati Shukla MNNIT Prayagraj, Allahabad, Uttar Pradesh, India

Anshy Singh Department of Computer Engineering and Application, GLA University, Mathura, India

Archana Singh Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

Juginder Pal Singh Department of Computer Engineering & Applications, GLA University, Mathura, India

Krishanveer Singh Institute of Business Management, GLA University, Mathura, India

Nishtha Singh Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India

Vineet Kr. Singh MUIT, Lucknow, India

Vivek Singh Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India

M. Sri Geetha Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India

R. Stephen Kristu Jayanti College (Autonomous), Bengaluru, India

P. Subha Department of IT, Sri Sai Ram Institute of Technology, Chennai, Tamil Nadu, India

M. Sujay Sudharshan Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India

S. Sunderesh Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India

G. Tamilmani Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai, Tamil Nadu, India

Pham Thi Bich Thao Thai Nguyen University of Sciences, Thai Nguyen, Vietnam

Pooja Tripathi Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India

Munesh Chandra Trivedi Department of CSE, NIT Agartala, IET, Agartala, India

Nguyen Dinh Trung National Economics University (NEU), Hanoi, Vietnam

V. Ulagamuthalvi Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

S. Uma Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai, Tamilnadu, India

R. Umamaheswari Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai, Tamilnadu, India

C. Vaithilingam Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India

Nguyen Van Thuy Banking Academy, Hanoi, Vietnam

Pham Van Tuan National Economics University, Hanoi, Vietnam

Neeraj Varshney Department of Computer Engineering and Applications, GLA University, Mathura, India

V. Vijay Easwari Engineering College, Chennai, India

A. Dalvin Vinoth Kumar Assistant Professor, Department of Computer Applications, School of Engineering, Dayananda Sagar University, Bangalore, India

Toan Linh Vu Le Faculty of Banking and Finance, Van Lang University, Ho Chi Minh City, Vietnam

Rohan Yadav Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India

Vikash Yadav Department of Technical Education, Lucknow, Uttar Pradesh, India

Amazon Alexa and Its Challenges to Reach More Households

Pankaj Pathak, Rati Shukla, Himani Jain, Vikash Yadav, Parashu Ram Pal, and Rishabh

Abstract Voice assistants became an important feature in the current smart device world. Taking instructions from human and providing services to them has increased the data traffic in the IOT systems due to which telecom players are getting benefitted. However, these devices are connected to the network all the time and without proper security measures will cause security breaches in the voice assistant systems like Alexa, Siri, Google Now, Cortana, etc. Alexa Amazon's smart speakers enabled with AI rapidly is adopted by households and inculcated in various daily life activities. Though the technology is innovative but with its dark side and its challenges of security vulnerability. The case study highlights major cybersecurity attacks on the voice assistants and the challenges which Alexa and other voice assistant devices are facing for their market expansion. This study enables to critique the security vulnerability in Amazon Alexa AI voice assistant and other such products. It helps to have profound understanding of key market and industry-based challenges in the current markets.

P. Pathak
Symbiosis Institute of Digital and Telecom Management Symbiosis International (Deemed University), Pune, India
e-mail: pankajpathak@sidtm.edu.in

R. Shukla
MNNIT Prayagraj, Allahabad, Uttar Pradesh, India

H. Jain
ABES Engineering College, Ghaziabad, Uttar Pradesh, India
e-mail: himani.jain@abes.ac.in

V. Yadav (✉)
Department of Technical Education, Lucknow, Uttar Pradesh, India

P. R. Pal
SAGE University, Bhopal, Madhya Pradesh, India

Rishabh
Galgotias College of Engineering and Technology, Greater Noida, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_1

Keywords Alexa · Voice assistant · False authentication · Dolphin attack · AI · Amazon

1 Introduction

Amazon Alexa is a well-known digital voice assistant, enabled with AI and used as smart speakers launched in 2014 by Amazon. Various enhanced versions came into market till now and reached to large households for various tasks. Alexa can be engaged in our daily life routine works, viz. controlling the home appliances, getting infotainment, scheduling our daily tasks, shopping, and many more. Alexa can play an effective role in home automation and to control various smart devices and merely by using voice. The Alexa system consists of Echo products identified as—the hardware and software that directly communicate with consumers and the cloud modules—that have most of the "smarts": automatic speech recognition, understanding natural language, and response. The working architecture of Alexa has been shown in Fig. 1. Third-party services through applications can also provide some responses via "skills." The third parties who write and publish those skills are responsible for the behavior of their skill. To demonstrate the working of Alexa, we are presenting an example to obtain weather information by simply requesting Alexa by speaking "Alexa, what is the weather." We can see how request flows which are caught by Echo and how it is interpreted, acted, and finally responded by Alexa. Amazon Alexa works on natural language processing which records the words and interprets the sound by using computational power. It then identifies words with matching in the database, makes sense of the tasks, and carries out corresponding function.

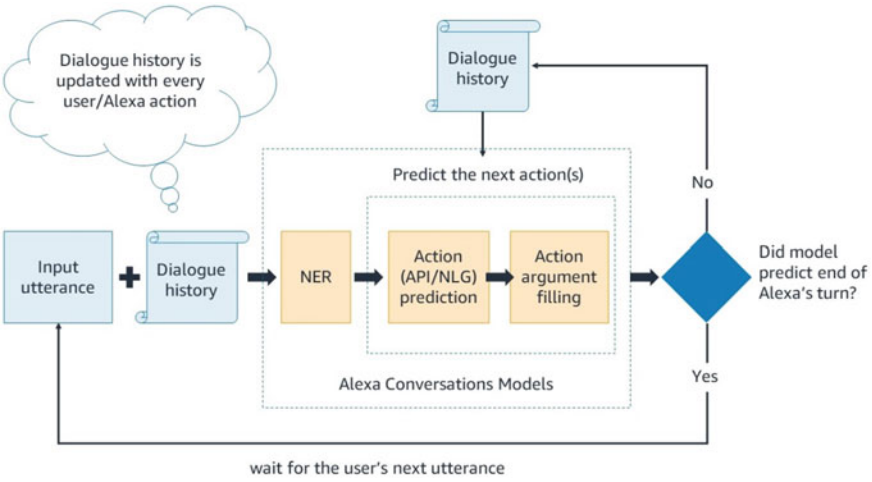


Fig. 1 Alexa working architecture [14]

Amazon Alexa was launched in 2014, and it is available for general public in 2015. Alexa's API availability to developers helps in expanding its market due to integrating it with non-Amazon products. This feature is amazing and exciting other brands to inculcate Alexa in their products. But at the same time, it also questions the quality of these third-party products. It may impact on the brand image of Alexa. But experts say that unless and until there are some products with good quality available there is no adverse impact on Alexa's brand image.

Why is Amazon Alexa so important? Alexa is successful in setting up a benchmark up to some extent for inaction machine learning with artificial intelligence as an effective consumer-based implementations.

Systems such as Alexa are an early sign of a significant shift of communication between humans and computers.

Who has affected by Amazon Alexa? Customers are primarily affected by Amazon Alexa who are using it in terms of positive as well as negative aspects. If we say positive, they ease their life by exploiting the services of it. If we see the dark side of it, then definitely security concerns and data ownership are the primary anxieties of it. Alexa also affects its rivals like Microsoft, Google, and Apple; all have digital assistants pushing Amazon to play catch-up.

After discovering that Amazon stores your audio recordings for Amazon employees to review, Alexa users grew even more alarmed. All this has made people a little bit cautious of using Alexa, and it is facing a major challenge of reaching different households. These privacy concerns are keeping people hesitating from using Amazon Alexa.

2 Evolution of Alexa

With the enhancement of digitalization, the market is seen to be demanding a more convenient and on the touch lifestyle. Amazon Alexa is one such digitally revolutionized AI-based product introduced by Amazon. As per January 2019, over 100 million Alexa devices have been sold [1].

Initially, Alexa was introduced with few utilities in the market but with the time these utilities are increasing. Alexa can perform various daily life tasks such as setting alarm, informing about weather, preparing list, scheduling task, accessing informative articles, etc. One of the interesting skills of Alexa is to listen live sporting events, National Geographic quiz, big sky as well as news stations [2]. Alexa also has built in support for Spotify and Pandora accounts.

As far as concern to technological advancements, in April 2019, Alexa-enabled devices could support the download of over 90,000 functions which was merely 1000 functions in 2016 [3]. The home automation feature was launched in April 8, 2015 [4]. Developers are able to create their own smart home skills using the Alexa Skills Kit. Alexa also emerged into the field of food delivery. Starbucks used Alexa for

placing pick up orders, but very few Alexa owners used these services for placing the orders [5, 6].

According to research conducted by Consumer Intelligence Research Partners (CIRP), in 2018, 69% of the US smart speaker market is captured by Amazon Alexa. And it is increased by 1% in 2019. As far as concern to virtual market, Alexa and the other competitors like Siri, Google Assistant, and Cortana had their market shares—25%, 36%, and 17%, respectively [7].

However, we see due to poor digital knowledge in India as well as a few other countries, Amazon Alexa is finding it difficult to reach out to the majority of the crowd. As of November 2018, Alexa is available in 41 countries. Also, Amazon Alexa is seen to have been facing strong competitions over the years which is one of the major hurdles faced by Amazon Alexa.

3 Challenges to Retain Household and New Market Expansion

3.1 Investigation of Security Vulnerabilities in Voice Assistants like Alexa

The reason for selecting these HDVA devices like Amazon Alexa is manifold. Amazon Alexa is flagship device in HDVA devices. There are 5 million devices being sold since its launch in 2014 within 2 years. Secondly, there are 10,000 skills (Alexa voice services) available for Amazon Alexa. Moreover, every smart device is equipped with voice assistant system, right from smart speaker to connected cars. Hence, exploring the security aspects of these devices will be helpful in finding out problems and advancement of these HDVA devices. Currently, Amazon Alexa takes voice commands from humans and performs some action. The actions include playing music, shopping online, checking weather, scheduling appointments, making payments, and controlling smart devices like garage doors, plugs, and thermostats. Basically, these HDVA devices have a feature like (always listening mechanism) which helps users to talk to the devices continuously so that users need not hold some button and pass commands but this will expose them to security vulnerability. In past there were several incidents came across which breached the security agreements.

3.2 Design of Voice Support Systems

Before understanding the security breaches of Alexa, let us understand the voice service model of Alexa. To control a smart device, a user can speak voice command to Alexa after waking up with a voice called “Alexa.” Alexa then sends a command to the remote cloud for authentication purpose using Wi-Fi. Once the command

gets authenticated, then the command is forwarded to the server called smart home skill adapter, which is maintained by Alexa in cooperation with third party. Later, the command is sent to another cloud which is used to control the corresponding devices.

3.3 Security Vulnerabilities in Alexa—False Authentication

The voice service facility in Alexa enables it to listen to users' commands [8]. However, it can respond and act to the commands during the absence of humans which is a security breach. When Alexa was designed, its design considerations were limited within the room and only home members can access it, but there can be some exceptions like the sounds accepted by Alexa can come from outside surroundings also which can lead to the compromising of security. Since there is no second factor authentication in these devices, it will lead to trust issues. After careful study, we observed that these kinds of issues are occurring when victims are not at home. If HDVA devices stop taking commands when users are not in the surroundings, then the fraudulent commands can be restricted. Alexa works on the scenario that when a user stands nearby, it takes commands from him/her and provides services. Now, we will exploit this feature by sending commands to Alexa using a Bluetooth speaker. We have used a Bluetooth speaker controlled by a smartphone at significant distance (within 8 m) from the Alexa. Interestingly, Alexa responded to such commands from the speaker.

Validation: Even though when a false command is passed by an unknown person to Alexa, it can be overcome by the Alexa authentication system, but the authentication system can be easily bypassed by generating commands using text to sound (TTS). There are many Web sites present that can be used to create mp3 sound of a speech; thereby, any device (mobile, Bluetooth, laptop) can be used to play audio and get access to Alexa.

3.4 Dolphin Attack Through Inaudible Voice Commands

Till now, we have seen cases where the HDVA devices can be activated without the human presence. In this regard, another attack has been explored on HDVA devices called Dolphin attack. Dolphin attack is basically injecting a sequence of inaudible voice commands (modulating human voice on ultrasonic carriers) that can lead to unnoticed security breaches to the voice-controlled systems. This is a technique to exploit ultrasonic sound channel (ultrasonic channel is the frequency channel which is greater than 20 kHz). Human's voice system is designed to listen voices up to 20 kHz. Mobile phones adopt audio sampling techniques lower than 44 kHz in which they deploy low-pass filter so that any signal greater than 20 kHz will

not be passed through it. Even though ultrasonic sounds are received and correctly sampled, they cannot pass through the sound recognition system of HDVA as they do not belong to the human tone. Dolphin attack overcomes all these issues and performs attacks. Dolphin attack may be imposed on Google, Siri, Cortana, and Alexa through sequence of inaudible voice commands. But all the hurdles like accessing malicious Web sites, spying, getting fake information, denial of service, concealing the attacks, listening the inaudible voice (>20 kHz), greater intelligence of inaudible sounds than speech recognition system, etc., can be overcome by dolphin attack and activate audio hardware of the device.

Any voice assistants will be activated using 2 commands: 1. activation commands like "Alexa," "hey Siri," etc., and 2. recognition commands like "Call 123456," etc.; Alexa uses speaker-independent algorithm that means it accepts "Alexa" word spoken by any individual as long as the word is clear. Whereas Siri gets trained only by a human, viz., it uses speaker-dependent algorithm. The speaker-dependent algorithm works on a local server, whereas the speaker-independent algorithm works on cloud server.

3.4.1 Attack Design

Dolphin attack is carried in 3 steps

1. Voice command generation
2. Voice command modulation
3. Voice command transmission

Voice command generation: The voice command generation is basically generating the activation words. Creating activation commands is different unless the user speaks "Siri." Here, the experiment will generate activation commands by two methods: 1. Attacker cannot find owner (stolen cases). 2. Attacker can obtain few recordings of owner. When attacker cannot find owner, TTS-based brute force technique is used which means we get recordings of human voices from TTS Web sites and then use them for activating. There are different TTS sources that provide human voices like Selvy Speech, Baidu, Neo speech, acapella, etc. Also the case when attacker obtained few recordings of victim. In this case we will use concatenative synthesis technique to extract the activation words. Let us say that the attacker got the recordings of victim like "City and Carry." So, sampling can be done to extract the word Siri like CIty + ca RRY = Siri. So, like this we can get activation commands in the first step.

Voice command modulation: Here, the malicious commands can be modulated over suitable carriers, while modulating select carriers of ultrasonic range so that they are inaudible. Then, amplitude modulation of baseband signal and ultrasonic signal has been obtained. The amplitude modulation parameters are like 1. modulation dept., 2. carrier frequency, 3. voice selection. Firstly, the modulation depth is set to 0.5 here which means the carrier amplitude varies 50% above and below its unmodulated level. Secondly, the carrier frequency is depending on many factors like frequency range

of ultrasounds, the cutoff frequency of low-pass filter, and the frequency response of the microphone. The lowest modulated frequency has to be 20 kHz to ensure inaudibility.

To impose successful attack for base signal of 6 kHz carrier signals, frequency must be of 26 kHz to ensure the lowest frequency is larger than 20 kHz. Also, different voices vary in terms of baseband frequency ranges. Let us say female voices have wideband than male voices due to which there is a probability of frequency leakage. So, it is preferred to select narrowband voices for the attack.

Voice command transmission: The voice command transmission consists of signal source, modulator, and speaker. Powerful transmitter is used with signal generator and portable transmitter with a smartphone. The first one is we are used to validate the dolphin attack and the second one is for walk-by purpose. The powerful transmitter uses all powerful equipment to generate, modulate, and send. The portable transmitter uses a smartphone to transmit the modulated signals. In the smartphones, we have low-pass filter that attenuates the higher frequencies in order to overcome these problems; narrowband ultrasonic transducers are used as speakers and amplifier to amplify them. In order to find how much the TTS-generated voice differs from a voice played by a smartphone, a method called Mel Cepstral dispersion may be used. It is observed that the MCD factor is less than 8 and hence is preferred for the attack.

Recognition versus activation: Various devices exhibit different results in terms of attack distance considered. In addition, if we append the controlling command to the recognition command, the chance of success is higher due to the fact that the activation commands are trained by speech recognition systems that are always on mode. The length of command matters when it comes to the activation of voice assistants. For example, "Call 12345/open abc.com" is harder to recognize than "Turn airplane mode" or "What's new today." In first case the word "open" and "call", has to correctly recognize for execution and in the second case only "turn" and "What" has to be recognized for execution. So, the attack can be successful if the attack word is short and commonly recognized by speech recognition system.

4 Implementation Issues

With the rise of voice-controlled AI-led voice assistants in the ICT ecosystem, there are still serious issues regarding the inhibitions about the same. Organizations are also looking into the business aspect of this technology and are considering an AI-led voice assistant ecosystem into their enterprise management to increase the efficiency and reduce the turnaround time (TAT) of the system. We will now look at some of the challenges [9, 10] for Amazon Alexa and its rivals which pose to business organizations for implementing the same.

- **Against Popular Opinion**

Fig. 2 Consumer Intelligence series voice assistant survey, 2018 [13]

It is a well-established fact among the general workforce that whenever anything automated is introduced into the system, there is always a fear of loss of jobs among the masses. Voice assistants are no exception. There is widespread fear that the introduction of AI will eventually cut our jobs.

- **Knowing the Real Areas of Action**

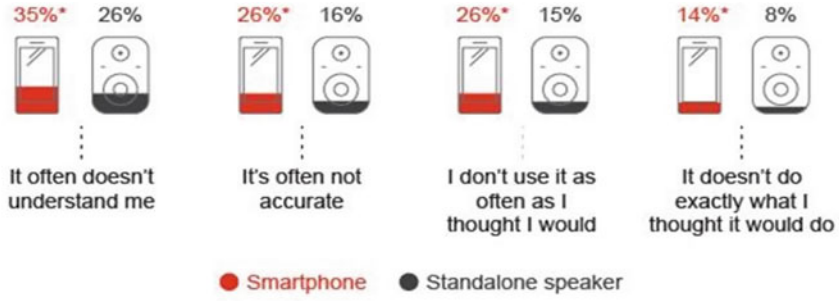
Even after your organization has finalized to go with an AI-led voice assistant (VA), the real challenge becomes to choose which vertical to implement and direct your R&D department to work upon and increase your productivity and efficiency. The verticals vary from HR, finance, technology, etc. Since there is very little information available about any known use cases of AI-led AI, it becomes a challenge for organizations to come up with one. A consumer survey was conducted in 2018 to check the usefulness of voice assistant in human's life as shown in Fig. 2.

- **On boarding and Integration of your Voice Assistant into the System**

Just like with any integration, your voice assistant (VA) will also have to be on-boarded and integrated into the system which is not just loading them with information or giving them access to all your resources. To be really efficient, your AI-led VA must know your entire flow-of-work as well as access to your databases and resources. This is a continuous and gradual task and does not happen overnight. Moreover, the real challenge [10] is to determine how much access you are going to give your VA so that you are not putting your organization at a business risk. Figure 3 shows a survey conducted among several consumers in 2018 to check the satisfactory rate of voice assistant device.

- **Handling Data**

The major unique selling proposition of an AI-led voice assistant is to process and analyze huge chunks of data and store as well as to make sense out of this data for later use. The main challenge is what to do with this huge database of data and how your



Q. Here are some reasons people claimed to be unsatisfied with voice assistants. Please select which one(s) apply to you. *Statistically significant at a 90% confidence level
Source: PwC Consumer Intelligence Series voice assistants survey, 2018

Fig. 3 Consumer Intelligence series voice assistant survey, 2018 [13]

AI engine handles this data without being a security risk. There is no standardized data handling framework existing as of now and are puzzled organizations to use voice assistants [8].

Amazon Alexa has practically 59% of the piece of pie in India followed by Google Home to 39%. Yet, it is very much deprived of arriving at each house in India. As indicated by the International Data Corporation (IDC)'s India Monthly Smart Speaker Device Tracker, a sum of 753,000 units was dispatched in 2018 in India. This number is very less if we consider the total Indian population. The consumer segment added to most of the keen speaker deals in the nation. In any case, the acquisition of brilliant speakers by ventures likewise saw a flood during the celebration or festive time, basically for gifting purposes, while the real endeavor reception for vertical explicit use cases stays at an incipient stage. With regard to the favored channel for savvy speakers, online channels comprehensive of e-trailers and merchant's own sites commanded brilliant speaker deals with 55% of the shrewd speakers sold in this channel. Current Web-based business infiltration in India stands at 28%, so we can clearly estimate the room for improvement.

At a public interview in New Delhi, the Web-based business mammoth said Alexa now coming with Hindi, a language verbally expressed by generally a large portion of a billion people in India, as the organization hopes to grow its compass in the country. To bring support for Hindi in Alexa, the company was working on it over a year. Users can now be able to ask Alexa their inquiries in Hindi, and the digital assistant will have the option to react in a similar language. The component, which will start turning out through a product update to Alexa gadgets beginning today, as of now just backings one voice type for Hindi. (For English, Alexa offers numerous voice types.) In the months to come, Amazon said it intends to include support for multilingual families, which will empower individuals from the family to associate with Alexa in the language they each like [9].

In view of its most current review, the firm said: "the US introduced base of smart speaker gadgets is 76 million units, up from 70 million units in the March 2019 quarter and 50 million units in the June 2018 quarter." Other studies contend there are in excess of 100 million units in American homes [11]. So, from these stats we can easily estimate that America being the less populated country is getting a huge number of Alexa shipped than India almost 10 times more than India.

5 Evaluation and Conclusion

5.1 Defense Mechanisms

The primary solution is concerns about today's microphones' sensing capability which can sense acoustic sounds, i.e., >20 kHz. Thus, a microphone has to be designed in such a way that it should suppress frequencies that are of ultrasonic range. This is called microphone enhancement.

Secondly, we can add a module before low-pass filter (LPF) to detect modulated baseband signal and demodulate them to obtain baseband signal. This process does not spoil the normal operation of microphone since there is no relation between captured voice signals and noises of ultrasonic range. This is called inaudible voice command cancelation technique. The voice recognition should not be passed through a random voice command.

Thirdly, we can use software-based defense mechanisms. This feature uniquely identifies modulated voice commands to the genuine ones. In this, we take a demodulated voice and genuine voice which both indicate a difference of 500–1000 Hz that means if we detect a signal in the range of 500–1000 Hz then we can detect the dolphin attack. To be specific, a machine learning-based model shall classify the voices. This method is called support vector-based technique which can be used as a defense mechanism against dolphin attacks.

Support vector machine (SVM) is a supervised machine learning algorithm mostly used in classification issues in which each element of data is plotted as a point in n-dimensional space (where n is the number of features you have) with the value of each function being the value of a particular coordinate.

5.2 Discussion and Analysis

In [12], authors discussed dolphin attack and they also presented hardware and software defense solutions. Our paper presented and discussed the challenges for Alexa which might be hurdle to get expansion in the market for voice assistant products. The study also discusses the impact of different attacks and the solutions presented for them. The research study enables to critique the security vulnerability in Amazon

Alexa AI voice assistant and other similar product which helps to have a profound understanding of key market and industry-based challenges in the current markets.

6 Summary and Recommendations

Voice assistants are becoming an important feature in the current smart device world. Taking instructions from human and providing services to them has increased the data traffic in the IOT systems due to which telecom players are getting benefitted due to heavy traffic. However, these devices are connected to the network all the time and without proper security measures will cause security breaches in the voice assistant systems like Alexa, Siri, Google Now, Cortana, etc.; hence in this case study, we have highlighted major cybersecurity attacks on the voice assistants. First one is false authentication in which the Alexa system is activated without sending human voice commands rather than sending the same through a recorded voice which indicates that a voice assistant system can be activated without the presence of humans. Secondly, we have discussed dolphin attack in which the ultrasonic frequency range is exploited by VA which is inaudible to humans. The modulation process of the ultrasonic sound on a high frequency carrier wave is discussed which can be used as an attack channel on victims' voice assistant system. We have also discussed how baseband signal can be generated in the form of several TTS systems and synthesis techniques, thereby generating activation commands from them. Several methods have been discussed, viz., how to prevent such attacks by modification of a voice assistant system like support vector machine algorithm which is a software-based firewall that can avoid dolphin attack and other physical changes of the system. We did analysis of voice assistant markets in the current position and the challenges like natural language processing, unable to process commands at every time, AI-based challenges, knowing the real areas of action, and handling data which became latest market challenges now.

Many other competitors in the market are into this same segment of the voice assistant. As discussed above in the case study, only 753,000 units of Alexa sold in the year 2018 which is very small concerning the total population of India. Here, the biggest challenge comes that how to increase this number in the coming years. Alexa should look at how to modify itself with the growing technology, and only voice assistant service is not enough to stand out in front of the competitors. Alexa can be integrated to control the IOT-enabled devices with the help of voice assistants run by mobile data provided by telecom operator, and it can increase the mobility of Alexa. This feature will make Alexa a more modified device which the Indian consumer can find worth buying.

References

1. Bohn D (2019) Amazon Says 100 million Alexa devices have been sold—what's Next? The Verge, 4 January 2019
2. Erickson S (2018) "Alexa, Make Me Money": conversational AI prepares for the enterprise, 22 September 2018
3. Press release (2019) Amazon.com announces first quarter sales up 17% to $59.7 Billion, 25 April 2019
4. Callaham J (2015) Amazon Echo owners can now control WeMo and Philips Hue devices with their voice, 8 April 2015
5. Kell J (2017) Starbucks adds voice ordering to iPhone, Amazon Alexa, 31 January 2017
6. Anand P (2018) The reality behind voice shopping hype, 6 August 2018
7. Olson C, Kemery K (2019) Voice report: consumer adoption of voice technology and digital assistants. Technical Report. Microsoft, 2019
8. Yadav V, Mishra D (2020) Home automation system using raspberry pi zero W. Int J Adv Intell Paradigms 16(2):216–226. https://doi.org/10.1504/IJAIP.2018.10017087
9. Healy P (2018) 9 challenges in implementing an AI virtual assistant into your organization, 30 April 2018
10. Chapman K (2018) The challenge of AI voice assistants in customer service, 7 June 2018
11. Sterling G (2019) Alexa devices maintain 70% market share in U.S. according to survey, October 2019
12. Zhang G, Yan C, Ji X, Zhang T, Zhang T, Xu W (2017) Dolphinattack: inaudible voice commands. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security, pp 103–117
13. PricewaterhouseCoopers (2018) Consumer intelligence series: prepare for the voice revolution. Retrieved from PricewaterhouseCoopers (PwC) Website: https://www.pwc.com/us/en/advisory-services/publications/consumerintelligence-series/pwc-voice-assistants.pdf
14. https://www.amazon.science/blog/science-innovations-power-alexa-conversations-dialogue-management

Fake News Detection Through ML and Deep Learning Approaches for Better Accuracy

Anil Kumar Dubey and Mala Saraswat

Abstract Social media form a very frequent podium to people to freely express their opinions and easily communicate to others. Nowadays, it plays a vital role for spreading the news headlines, and it became most applicable news sources globally as easily accessible, but also risky as exposure of "fake news" misleads the people. The extensive spread of such misinformation deploys negative impacts on people and society and becomes recently a global problem. Several issues already rise in worlds during elections process, due to huge spread of fake news. Therefore, the detection of it on social platform transforms into an emerging research that is exciting enormous concern. Problem to identifying the fake news has concentration to public as well as government organization. Such propaganda probably affects the opinion of people and malicious parties involved to manipulate the conclusion. Due to the majority of society opinion impact changes, fake news detection is an important challenge to researchers. The detection of misinformation is not an easy task for anyone, but quite is a complex for people. Here, we analyze the different fake news detection approaches followed in current scenario and compute the detection process through machine learning and deep leaning algorithms for better accuracy.

Keywords News · Social media · Detection · Classification · Accuracy

1 Introduction

Several misinformation spread on society through social media and lead to change the opinion of people. The detection of this distinguishes from fact in it and is very important task for society. Researchers of several areas are investigating the mounting production as well as diffusion of misinformation that rapidly infects the society. "The news that is purposely and certainly false and is able to mislead the readers" is known

A. K. Dubey (✉) · M. Saraswat
ABES Engineering College Ghaziabad, Ghaziabad, Uttar Pradesh 201009, India

M. Saraswat
e-mail: mala.saraswat@abes.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_2

as fake news [1]. According to Merriam Webster Online Dictionary, fake news is "the news reports that are intentionally false or misleading" [2]. Generally, we can say that fake news is unauthenticated information that circulates or propagated on any platform for misleading to human brain, and it includes rumors, propaganda, and satire [3].

Basic

The mounting curiosity to detect the fake news fascinated to researchers due to the circulation of fake information through social media platforms. Such misinforming content is quickly spreads on social media and gains popularity. As people received the content from social media, they easily can believe on it and interpret their minds for it as reliable source of information to get trust. Due to blind trust and easily acceptance of people for fake news information, several solemn and unenthusiastic impacted fake articles get viral and observe on society and people leads to a disparity of news ecosystem [4]. Nowadays, social media (i.e., Facebook, WhatsApp, and Twitter) have covered most of the developing countries worldwide to become the major resource for news. It is essential in economics, social developments, and politics for motivated human brain and influences the process of these in negative impact and finally targets to damage to public figures and agencies [5].

Several active researchers are enduring to detect the fake news from social media. The detection process is estimating the misleading information of any news/topic whether it is planned or accidentally affected to society. In most of the cases, the detection process deploys the machine learning algorithm to filter the news, whether it is fake or not [5]. The detection of such things is very difficult task due to slight differences between fake and real news [6]. Current scenario of messaging shows that the fake news has become an ordinary thing for it. No one can refute the contents easily as it may fake; still their meaning depends on view of person [7]. The identification of fake news is an imperative subject for both public and society. The amplifying social media interaction has led the increment in number of people to unfair [8].

Problem

From past few decades, social media cover globe as they become the major resource of news information, due to the accessibility, minimum cost. Other perspective illustrates the danger to expose "fake news" aimed to mislead and manipulate the readers. The extensive spread of such misinformation deploys negative impacts on people and society and becomes recently a global problem. One major issue rises in U.S 2016 presidential elections due to huge spread of fake news. Therefore, the detection of it on social platform transforms into an emerging research that is exciting enormous concern.

Objective

The main objective of this work is accurate detection of fake news using machine learning algorithms.

2 Related Work

Researchers' enduring valuable contribution in this era is to distinguish the fake news from facts. Kaur et al. study and evaluate the superior supervised algorithm to detect the fake news. Author concerns about to superior learning classifier of machine learning on the basis of diverse conditions for detection, and appropriate model that gets best detection in particular condition is elaborated. Due to fake news getting viral on social media related to Covid-19 will impact society because our prediction will be based on data set collected from fake viral news. Similarly, a fake video was spread few years before on social media related to Kerala battling for floods, and it became so much viral. The news claims that the Chief Minister of Kerala state was forced to Indian Army to stop the rescue operations in state of flooded regions. One of the famous fake news was viral through WhatsApp groups in India in 2019 during national election was impacted to India's ruling party [4]. Harita Reddy analyzed several approaches to detect the fake news through text features. Author gets the 95.49% accuracy in detection through combo of stylometric and text features [9]. Nicollas et al. elaborate the analysis to detect the fake news through text extraction of social media, using the natural language processing. Author uses news data from Twitter as 33,000 tweet and distinguished the real and fake news among them. Approx. 86% accuracy is received through dimension reduction of original features [10]. Rubin et al. discuss about the details of fake news as stated that it can be divided into three parts as: pure fraudulent nature target to confuse the readers, rumors, and sarcasm and irony [11].

According to Peining Shi et al., malicious social bots are spreading the misinformation to mislead the society, therefore, wish to detect and remove these bots from social networks. Generally, this detection uses the easily imitated quantitative features for behavioral analysis and receives the low accuracy of detection. Author presents the joint approach as transition probability-based feature selection and semi-supervised clustering for detection [12]. Ghafari et al. discuss the trust concept for social networks and trust-related challenges to prediction process. Author classifies the trust prediction through addressing the challenges and invites the contributor for this era [13]. Day to day the emerging technologies arise, a need of viral reduction methodology is acquired for fake news to control the misleading of society. Shrivastava et al. present a model to evaluate the fake news propagation and describe how fake news spreads among several groups. Authors considers the current pandemic as COVID-19 for viral fake news [14]. Umer et al. discuss a fake news detection stance model based on headlines and news body. Author used the principal component analysis, chi-square for quality features extraction and also concerns to dimensional reduction approach for better result. PCA is used for noise removal and discusses model gain approximately 97.8% accuracy [15].

Domenico and Visentin discuss the marketing-related fake news and studied their details as consumer behavior, marketing ethics, future avenue, and strategy for fake news from eighty-six scientific articles and five managerial reports [1]. Ajao et al. present the fake news characteristics related to sentiments and process of

fake news detection. Author analyzes the text-based fake news detection considering both included and excused sentiments on Twitter dataset [3]. Elhadad elaborate the systematic survey of fake news detection on social media till 2017. Author discussed different types of fake news and presents general overview of summarization of news documents with different features that are extracted from news. Author notices that as spreads of fake news in social media, the detection system is not sufficient and its shortage invites researchers for more contribution in this era. Several prospectives are vacant for detailed contribution in big data of fake news [5]. Kuai Xu et al. highlight the continuous growth in fake news on social media that impacted the society. Authors target to analyze the differences between real and fake news based on their status and domain uniqueness. Kuai et al. used neural network for distinguishing the text in high-dimensional vector space for analysis [6]. Wenlin Han and Varshil Mehta discuss and evaluate the performance to detect the fake news in social networks through machine learning as well as deep learning algorithm. The fake news spreads rapidly in society leading to misguide the opinion of people, due to the fasted and easiest medium to transmit the information. The misleading information creates major impact on reader's brain for manipulated aspects. Authors use naïve Bayes, hybrid convocational neural network, recurrent neural network algorithm for it [16]. Hanz and Kingsland discuss about a news that is real or fake in details. As in presidential election 2016 at the USA, it created lot of information to mislead the people and impacted their brain. One workshop was organized to discuss the hole and flaw of viral information for election and analyzed the tweets to compute the reality and compared from previous [17]. Rajesh et al. discussed about a classifier to predict reality in viral news slice. Authors used the several years' news headlines to compare and for prediction process of news reality through natural language processing to mine the text [7]. Correia et al. focused to detect the fake news with new feature extraction, and analysis for practical application, and also concern about offers as well as challenges for it [18]. Day to day, the fake news identification becomes most popular issues for society, due to growth of social network users. Vereshchaka et al. stated that the fake news becomes the issues of not only for individuals but also as societal issues due to continuous people growing interaction on social media and technical challenges to distinguish the fake and reality of news. As per statistics of research, more than two millions of users deleted every month by famous social media as WhatsApp to stop the spread of the misleading information [8].

As several researchers already contributed in this era of fake news detection, but still some more efforts are required for detection as day to day grow-up in the social media users, so researcher is continuing to work in this era for more accurate and advance detection of fake news.

3 Proposed Approach

We collect data from Kaggle and preprocess the data for missing and unwanted data. After preprocessing, the different ML algorithms will perform one by one and check the accuracy of fake news detection. The algorithm which gains higher accuracy is pointed out. Decision tree algorithm and XGBoost provide the best accuracy in prediction of fake news. Author also applied the long short-term memory (LSTM) algorithm to predict the higher accuracy for ideal condition of acceptance.

Data Preprocessing

Before applying the classification techniques of decision tree, it is required to preprocess the data, for a definite alteration as shuffle, stop word and punctuation removal from text, grouping, lower casing, word clouds, and tokenization. The preprocessing process optimizes the data as per requirements from original size. The general preprocessing techniques are used to remove punctuation and non-letter typescript; after that, the lowered casting is performed. In addition, word cloud is used to represent the words in graphical way, and tokenization is done to count number of tokenized data frame. Stop words are irrelevant words normally used in sentence for their structure formation and generate the noise during classification. These words are removed from original data, and processed data is stored for next step (Fig. 1).

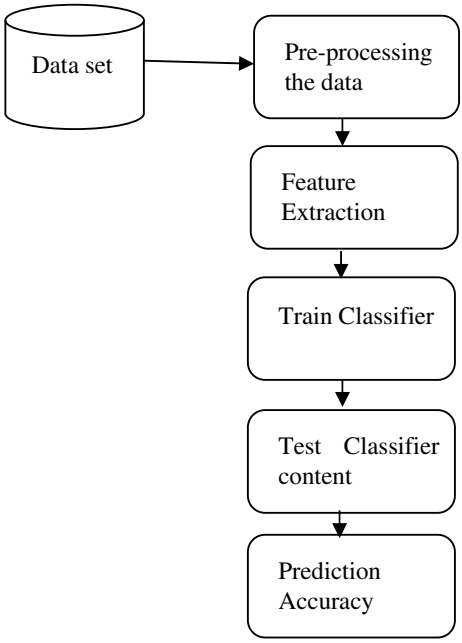


Fig. 1 Proposed approach for fake news detection

Features extraction

Several terms, phrases, and words may present in the data that show the extra load for computational to the learning process, and also some irrelevant features impact the classifier performance and accuracy. Therefore, its feature reduction is very important task that reduces the features size in feature space dimension.

Train Classifier

Select the appropriate classifier as decision tree for classification and split the data into two parts as training and testing. The target plan for classifier training is up to eighty percent of text data using random state.

Test Classifier

After training process of text data, the testing phase continues with target plan up to twenty percent of text data using random state. Prediction Accuracy. The decision tree is most popular technique for prediction and classification. The decision tree classifier accuracy for false news prediction will be computed with considering parameters.

4 Experiments

In this section, we demonstrate the detection through decision tree as best machine learning algorithm for prediction with Kaggle datasets for accuracy of detection. After that, we applied XGBoost algorithm and LSTM algorithm for better accuracy [19, 20]. In considered data, the shape is (44,898, 5) defined, in which (23,481, 4) fake news and (21,417, 4) true news shape ordered.

Figure 2 illustrates the graphical representation of word cloud data in database for fake and real news. The first part of figure as (a) represents the fake news world cloud; similarly, (b) represents true news world cloud. This operation performs through preprocessing of data, in which the world counted and word text-based cloud are formed. Decision tree is the supervised machine learning algorithm for continuously

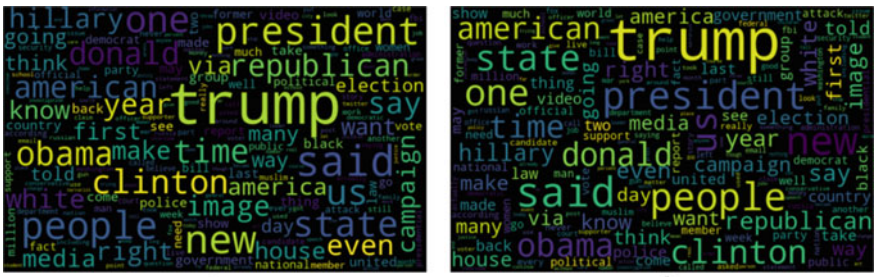


Fig. 2 **a** Fake word cloud, **b** true word cloud

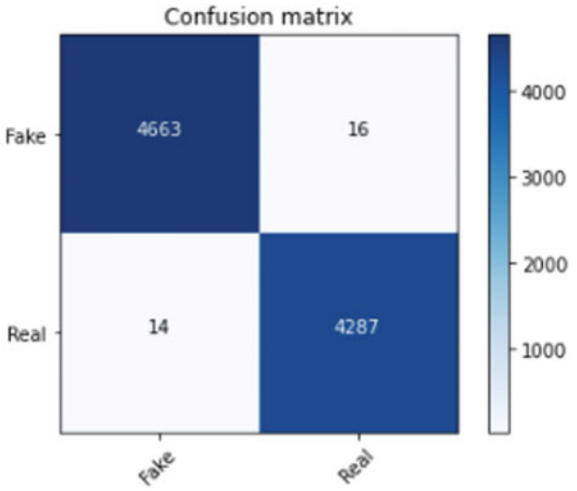


Fig. 3 Confusion matrix of decision tree for fake news detection

splitting data based on certain parameter. This classifier is used to divide and conquer approach to split data into subsets and again subsets as required. Therefore, author considers this algorithm for prediction of fake news. By applying the vectorizing the text in pipeline of vector count with maximum depth of tree considered as twenty and random state up to forty-two for transformer, gain of the confusion matrix is as illustrated in Fig. 3.

After applying the decision tree algorithm, ninety-nine point six seven (99.67%) percent for prediction of fake news is captures. After successfully applied the decision tree, author processes for best accuracy and applies decision tree-based ensemble ML algorithm (XGBoost) where a gradient boosting framework is used. Author is also applying the LSTM on dataset. LSTM is the type of recurrent neural network that has the capability to learn order dependency in sequence prediction problems. After processing both, author compared the result as mentioned in Table 1.

After applying the XGBoost algorithm, ninety-nine point seven (99.7%) percent for prediction of fake news is captures, that is, greater than the decision tree algorithm. Finally, the accuracy ninety-nine point nine (99.9%) percent for prediction of fake news is captures from long short-term memory. As it is near about to ideal condition of prediction as hundred percent, author did not check another algorithm for prediction.

Table 1 Comparison of performance of ML and deep learning approach for fake news detection

Method	Accuracy	Precision	Recall	F1-score
XGBoost	0.997	0.85	0.85	0.85
LSTM	0.999	0.87	0.87	0.87

5 Conclusion

The digital age of technology motivated people to interact with social media for news and messages. Due to high interaction of population of society, people post, transfer, and gain the news as well as messages from this. And, some illegal group is disturbing this phenomenon of accepting the news via posting the illegal information. As the human brain mostly faiths on it, it cannot distinguish the viral fake news and accept the viral news as real news and society as well as individual's brain is changed for that. Therefore, it is very important task for organization to control such rumors to spread from society and also detect the fake news. Author contributed in this era is targeting to accuracy for better prediction. As decision tree classifier is providing the best solution in most of the cases of prediction, therefore, applying this algorithm is to predict the fake news and gain most acceptable accuracy of prediction. After that, author also applied XGBoost and LSTM algorithms for better accuracy and received ideal condition of acceptance to accuracy. In the future, author plans to use different type of complex data and big data of fake news for classification to be capture with the best accuracy.

Conflict of Interest Statement On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Domenico GD, Visentin M (2020) Fake news or true lies? Reflections about problematic contents in marketing. Int J Mark Res 62(4):409–417. https://doi.org/10.1177/1470785320934719
2. Fake news—political scandal words. [Online]. Available https://www.merriam-webster.com/words-at-play/politicalscandal-words/fake-news
3. Ajao O, Bhowmik D, Zargari S (2019) Sentiment aware fake news detection on online social networks. In: ICASSP 2019. 978-1-5386-4658-8/18/$31.00 ©2019 IEEE, pp 2507–2511
4. Kaur S, Kumar P, Kumaraguru P (2019) Automating fake news detection system usingmulti-level voting model. SpringerVerlag GmbH Germany, part of Springer Nature 2019. https://doi.org/10.1007/s00500-019-04436-y
5. Elhadad MK, Li KF, Gebali F (2019) Fake news detection on social media: a systematic survey. 978-1-7281-2794-1/19/$31.00 ©2019 IEEE
6. Xu K, Wang F, Wang H, Yang B (2020) Detecting fake news over online social media via domain reputations and content understanding. Tsinghua Sci Technol 25(1):20–27. ISSN 1007-0214 03/14. https://doi.org/10.26599/TST.2018.9010139
7. Rajesh K, Kumar A, Kadu R (2019) Fraudulent news detection using machine learning approaches. In: 2019 global conference for advancement in technology, India. 978-1-7281-3694.3/19/$31.00 ©2019 IEEE
8. Vereshchaka A, Cosimini S, Dong W (2020) Analyzing and distinguishing fake and real news to mitigate the problem of disinformation. In: Computational and mathematical organization theory. S.I.: SBP-BRIMS 2019, © Springer Science+Business Media, LLC, part of Springer Nature 2020.https://doi.org/10.1007/s10588-020-09307-8

9. Reddy H, Raj N, Gala M, Basava A (2020) Text-mining-based Fake News Detection Using Ensemble Methods. IJAC, © Institute of Automation, Chinese Academy of Sciences and Springer-Verlag GmbH Germany, part of Springer Nature 2020. https://doi.org/10.1007/s11633-019-1216-5
10. de Oliveira NR, Medeiros DSV, Mattos DMF (2020) A sensitive stylistic approach to identify fake news on social networking. IEEE Sig Process Lett. https://doi.org/10.1109/LSP.2020.3008087
11. Rubin VL, Chen Y, Conroy NJ (2015) Deception detection for news: three types of fakes. In: 78th ASIS&T annual meeting: information science with impact: research in and for the community. American Society for Information Science, p 83
12. Shi P, Zhang Z, Kwang K, Choo R (2019) Detecting malicious social bots based on clickstream sequences. IEEE Access.https://doi.org/10.1109/ACCESS.2019.2901864
13. Ghafari SM, Beheshti A, Joshi A, Paris C, Mahmood A, Yakhchi S, Orgun MA (2020) A survey on trust prediction in online social networks. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3009445
14. Shrivastava G, Kumar P, Ojha RP, Srivastava PK, Mohan S, Srivastava G (2020) Defensive modeling of fake news through online social networks. IEEE Trans Comput Soc Sys. © IEEE 2020. https://doi.org/10.1109/TCSS.2020.3014135
15. Umer M, Imtiaz Z, Ullah S, Mehmood A, Choi GS, On BW (2016) Fake news stance detection using deep learning architecture (CNN-LSTM). IEEE Access. https://doi.org/10.1109/ACCESS.2017
16. Han W, Mehta V (2019) Fake news detection in social networks using machine learning and deep learning: performance evaluation. In: 2019 IEEE ICII. 978-1-7281-2977-8/19/$31.00 ©2019 IEEE. https://doi.org/10.1109/ICII.2019.00070
17. Hanz K, Kingsland ES (2020) Fake or for real? A fake news workshop. Ref Serv Rev 48(1):91–112. © Emerald Publishing Limited, 0090-7324. https://doi.org/10.1108/RSR-09-2019-0064
18. Reis JCS, Correia A, Murai F, Veloso A, Benevenuto F (2019) Supervised learning for fake news detection. Affective computing and sentiment analysis. IEEE Intell Syst. 1541-1672_2019 IEEE, Published by the IEEE Computer Society. https://doi.org/10.1109/MIS.2019.2899143
19. Sahoo SR, Gupta BB (2021) Multiple features based approach for automatic fake news detection on social networks using deep learning. Appl Soft Comput 100:106983
20. Choudhary M, Chouhan SS, Pilli ES, Vipparthi SK (2021) BerConvoNet: a deep learning framework for fake news classification. Appl Soft Comput 110:107614
21. Jindal R, Dahiya D, Sinha D, Garg A (2022) A study of machine learning techniques for fake news detection and suggestion of an ensemble model. In: International conference on innovative computing and communications. Springer, Singapore, pp 627–637
22. Sharma DK, Garg S (2021) IFND: a benchmark dataset for fake news detection. Complex Intell Syst 1–21
23. Zervopoulos A, Alvanou AG, Bezas K, Papamichail A, Maragoudakis M, Kermanidis K (2021) Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests. Neural Comput Appl 1–14
24. Kaliyar RK, Goswami A, Narang P (2021) FakeBERT: fake news detection in social media with a BERT-based deep learning approach. Multimedia Tools Appl 80(8):11765–11788
25. Khanam Z, Alwasel BN, Sirafi H, Rashid M (2021, March) Fake news detection using machine learning approaches. IOP Conf Ser Mater Sci Eng 1099(1):012040. IOP Publishing
26. Divya TV, Banik BG (2021) A walk through various paradigms for fake news detection on social media. In: Proceedings of international conference on computational intelligence and data engineering. Springer, Singapore, pp 173–183
27. Dubey AK, Singhal A, Gupta S (2020) Rumor detection system using machine learning. Int Res J Eng Technol (IRJET) 07(05). e-ISSN 2395-0056
28. Hakak S, Alazab M, Khan S, Gadekallu TR, Maddikunta PKR, Khan WZ (2021) An ensemble machine learning approach through effective feature extraction to classify fake news. Futur Gener Comput Syst 117:47–58

Blockchain Technology Concepts, Applications and Issues—A Survey

Amrita Jyoti, Rashmi Mishra, Vandana Kumari, and Shweta Chauhan

Abstract Blockchain is a growing technology that is particularly evident in the division of local power, persistence, anonymity, and auditing. The amount of data generated at these sites is huge, and the size of the blockchain is growing slowly as data transactions are maintained. To address this backwardness, blockchain technology needs to be translated, especially in relation to size. Certainly, afterward a limited age of process, the Bitcoin and Ethereum blockchains previously consider more than 200 GB, making it unbearable to accomplish partial devices. Several solutions have been suggested, each with a different approach. BC includes consensus, before the series was reviewed, the traditional procedure of proof of work (PoW) was used for consensus, we all know that it uses high power, and by using high computational power, it makes unsuitable for WSNs when given that they send data at a fast rate. Proof of work procedure is not able to handle this heavy load. Therefore, we can go to the advanced procedure of stakeholder evidence (PoS) or another popular consensus procedure such as Practical Byzatine Fault Tolerance (PBFT) and proof of authority (PoA). We also plan to adhere to a single (BS) station with multiple gateways that performs BC renewal after contract based on the material obtained from the nodes. Wireless sensor networks (WSNs) stereotypically encrypt thousands of sensors reserved for investigation, data collection, and communication on remote servers for unremitting procedure.

Keyword Blockchain

A. Jyoti (✉)
Ajay Kumar Garg Engineering College, Ghaziabad, U.P, India

V. Kumari · S. Chauhan
Lloyd Institute of Engineering and Technology, Gr. Noida, U.P, India
e-mail: vandana.gupta@lloydcollege.in

S. Chauhan
e-mail: shweta.chauhan@lloydcollege.in

R. Mishra
Krishna Engineering College, Mohan Nagar, Ghaziabad, U.P, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_3

1 Introduction

Blockchain technology with dispense structures provides a new resolution for the dispense WSN scheme. In this paper, a validation program based on the WSN verification blockchain is planned. WSN nodes are split up into base channels, cluster head nodes, and nodes, where nodes are divided into multi-level such as normal node, advance node, super node, super advance node, which are built into a sequential network. A blockchain network is built between a variety of nodes to form a hybrid blockchain prototypical, counting a local chain and a social network. In this hybrid model, authentication of the identity of the node in several announcement environments is visible, the normal functionality of the node authentication is achieved by the local blockchain, and authentication of the cluster head identity is visible in the public blockchain [1]. The safety and presentation analysis shows that the system has complete security and better presentation. Sensor nodes in wireless sensor networks (WSNs) are known for fast data transfer, so there is a need for authentication authority enabled for the function of secure and translucent migration of data from global locations across the network. We, therefore, need a blockchain (BC)-based network that includes transactions that can be duplicated and dispense over multiple notes to a peer-to-peer network that allows data reliability and flexibility. It is also protected from the network failure if nodes get compromised or die due to the heavy power feasting [2].

2 Blockchain

Blockchain is a data structure or ledger that stores information about any transactions that occurs (not just Bitcoin). Anything that is stored once cannot be changed or modified. This feature of blockchain makes it the most secure. It is decentralized and establishes a peer-to-peer network thus eliminating any middle-men.

Blockchain is made up of blocks that are stored in a chronological order. Each block has a capacity of around 500 transactions on an average. Thanks to the cryptography involved, these blocks are extremely secure [3]. Each block will a unique "hash" value attached to it which is calculated based on the data stored in the block. Every time a new block is added to the chain, the new block contains the hash of the previous block as well. So, modifying the contents of any previous block is practically impossible (and would destroy the whole chain). This makes blockchain immutable [4]. Some advantages of blockchain are.

2.1 Blockchain is Immutable

It means once information has been printed to a blockchain, no one, not even an organization manager, can modify in it. The blockchain can be transformed in attach only style. In other words, transactions can only be appending in the blockchain. Alterations and erasures are not permissible.

2.2 Blockchain is Decentralized

It means that it does not rely on a dominant theme of regulator. Everything is spread over. Those make the system fair and secure.

2.3 Smart Contracts

These are a set of protocols or more like a computer program that are stored in a blockchain and are executed when certain conditions are met.

2.4 Consensus Protocols

These are a set of protocols that keep all the nodes in the network synchronized with each other. It stops any solitary object from regulatory the whole blockchain scheme. The goal of consensus protocols is to assurance a solitary chain that is rummage-sale and shadowed.

2.5 Hashing

It is process in which the data stored in a block is converted into a fixed length output through a mathematical algorithm (e.g., SHA-256 is used by Bitcoin). The value of hash is unique for the same data, and it is impossible to produce the same hash using different pieces of data.

2.6 Wallet

A Bitcoin wallet is like a physical wallet. It contains your private key which can be used by you to store/spend Bitcoins on the blockchain.

2.7 Private Key

It is similar to the passwords we use to initiate a transaction. Likewise, using these private keys, we can spend Bitcoins from our wallets using a cryptographic signature.

Blockchain technologies concepts come to remove the centralized authority for the authentication process and add the transparency to everyone in the BC. Open permissioned BC typically follows to this value of transparency even nevertheless rather central in symbols its information and might be valuable for requests such as individuality schemes, hypothetical guarantee schemes, anywhere anybody can recite its information, but only a sure set of members is advantaged to inscribe the information into blockchain [5]. Closed permissioned BC is completely central and also not translucent to anybody, dismantlement the crucial perception of a blockchain. Consequently, these blockchains can be substituted with dispersed record schemes with limitations applied on the highest of it.

3 Digital Ledger

Blockchain is a moral digital ledger of financial communications that can be involuntary to record not just economic communications but practically the whole thing of worth. There are many definitions which are defined by many authors, such as blockchain is a type of data structures in which blocks are managed in chronological manner with the timestamp. Blockchain is the dispersed numerical ledger of an unchallengeable community greatest of numerical dealings [6, 7]. All novel greatest is authenticated crossways that the dispersed system beforehand it is stowed in a chunk. Every material once kept on the ledger is demonstrable and auditable but not editable. Each block is recognized by its cryptanalytic signature. The primary block of the blockchain is recognized as the genesis block. Figure 1 shows the digital ledger.

To access data of the first ever created block, one has to decussate from the last block created to the first block. The main feature of the blockchain is to overcome the challenges faced by the current system such as user has to pay the transaction in the present scheme (e.g., banks and other tertiary revelries take fee for moving cash); in the present scheme, the smallest applied contract scope is partial; wounding off the option for minor unpremeditated contacts, economic connections are unhurried and scheme is impervious and absences slide and justice. Also, the chief expert in

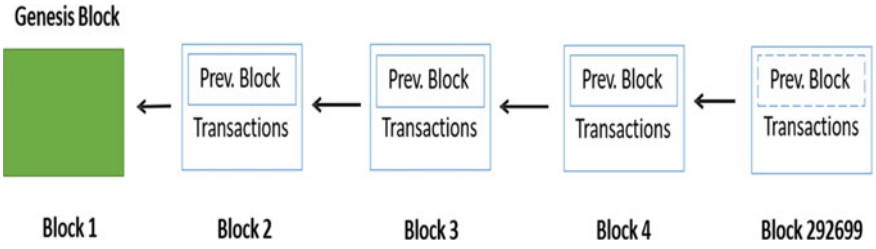


Fig. 1 Digital ledger

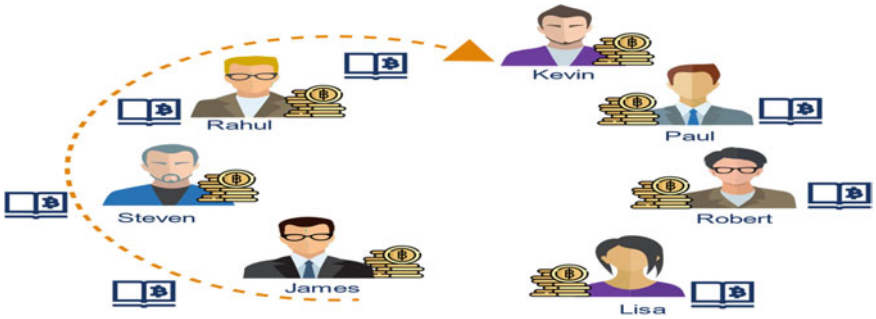


Fig. 2 Blockchain network

switch can overemployment the control and generate cash as per their own will. So, to overwhelmed the overhead contests, we essential such schemes which remove the essential of traders or tierce parties, thus creation contract charges nil or negligible, improves the contract implementation speed and can enable prompt settlement, is see-through and interfere hardy in command to circumvent operation or misappropriation, safeguards that money formation is not in the regulator of any dominant expert, is controlled to preserve the worth of the cash. We can say that the distributed system is the solution for all the problems [8, 9], which says the distributed system allows a system of mainframes to uphold cooperative accounting via the net. The distributed scheme is that this is exposed and is not in the regulator of any one gathering, and it is obtainable in one register which is entirely disseminated transversely the system. In distributed system, scheme where binary or additional nodes effort composed in a synchronized style to attain a mutual product. It is demonstrated in such a way that end operators see it as a solitary rational stage.

3.1 Working of Blockchain

As shown in Fig. 1, suppose Kevin wants to send the data to the Paul, then both have to be the part of the blockchain network. All the user needs to have public and

private key pair, and the user is identified in the network by using the public key. To view the transaction in the blockchain, user uses their private keys. Now, Kevin initiates the genesis block including the address of its own and the address of the receiver, timestamp, token, hash value, transaction information, etc. [10]. The block is validated by all the nodes in the network, then only block generated by the Kevin is added into the chain. Therefore, the block is received by the Paul. Paul decrypts the transaction using cryptographic algorithm (SHA-512). To ensure the security and the safety in a digital relationship, blockchain works in three main principles, such as distributed ledgers, authentication, public key cryptography [11].

Distributed Ledgers: In the distributed ledger, there is no central authority who maintains the ledger. All the updating is done in real time basis by all the nodes in the network. This process hardly takes few seconds to reflect the changes in the ledger.

Authentication: Figure 3 shows the authentication of transaction process in the blockchain. Before adding any block or transaction in the chain, the block needs to be authenticated by means of validated or verified by the users using some algorithm; therefore, this transaction is considered as the genuine. As soon as the transaction is validated by the user, it is encrypted, signed, and stored digitally. Contract is disseminated in the method of a digital communication. Just like your name and signature deliver the resistant of possession on the article, the digital signature affords the proof that the transaction is unaffected. Unlike a handwritten password, the digital signature is single for each business [12]. After the contract is fashioned, it broadcasts in a disseminated system.

Public Key Cryptography: Public key cryptography is also identified as asymmetric key cryptography. PKC uses the different key for the encryption and decryption process. The key pair is known as public key or private key. Private key is similar

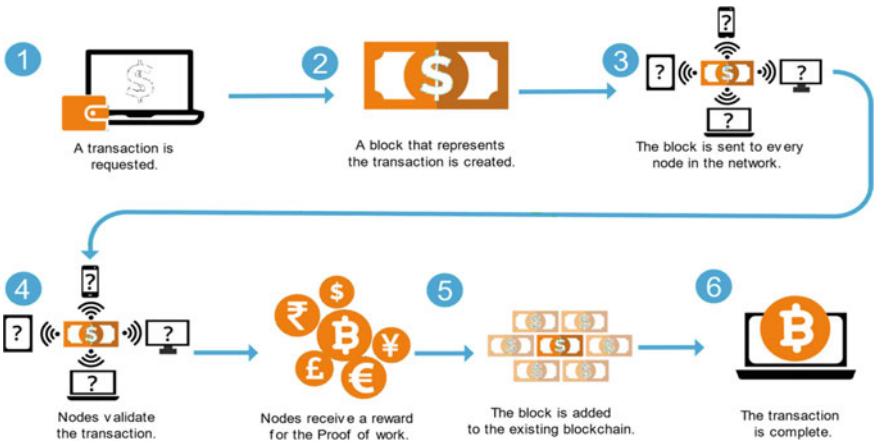


Fig. 3 Authentication of transaction

to the passwords we use to initiate a transaction. Likewise, using these private keys, we can spend Bitcoins from our wallets using a cryptographic signature [13, 14].

Currencies in Blockchain: With the implementation of Bitcoin, blockchain technology application was founding on decentralized currency. These currencies are known as cryptocurrencies and divided into coin and tokens. Token runs on the top of the blockchain, whereas the coins run separate on the blockchain. For example, Ethereum is a separate coin and a cryptocurrency which turns on Ethereum's blockchain where as there are many cryptocurrencies which run on the existing Ethereum blockchain, examples of these algorithms are ERC-20 tokens such as Binance Coin, Tether USD, Maker, and Basic Attention Token are cryptocurrencies [14]. Alike to fat moneys, cryptocurrencies likewise connected value originated on the trust people devour on the detailed cryptocurrency. Smooth though vicissitudes typically assistance the energy of a currency, tokens are characteristically representation of a benefit for a formation, capability, an advantage, or smooth a precise. Well-documented cryptocurrencies are open source, and novel coins are produced from a computational procedure distinct with an agreement instrument putative by its public starting the unchallengeable blockchain of the specific cryptocurrency [15]. These apparatuses safeguard the truthfulness and immutableness of the blockchain-based cryptocurrency ledger short of dominant authority regulatory its administration. Table 1 shows the types of cryptocurrencies used in blockchain.

Table 1 Types of cryptocurrencies in blockchain

Year	Cryptocurrency	Mining methods	Technology used
2008	Bitcoin	Proof of work	SHA-256
2011	Litecoin	Proof of work	Scrypt, Fork of Bitcoin
2012	Ripple	NA	Ripple
2013	Ethereum	Proof of work	Ethash
2014	Tether	NA	Issued by Tether Limited
2014	Steller	NA	Steller Development Foundation
2014	NEO	Proof of stake	Smart contracts to issue and manage digitized assets
2015	Cardano	Proof of stake	Ouroboros proof of stake
2017	EOS	NA	ERC-20, EOS mainnet
2017	Binance Coin	NA	ERC-20
2017	Bitcoin Cash	Proof of work	SHA-256, Fork of Bitcoin to increase block size
2018	Bitcoin SV	Proof of work	SHA-256, Fork of Bitcoin cash with different block size

4 Applications of Blockchain

Blockchain was familiarized with Bitcoin whitepaper to determine the double-spending problematic of electric money in a dispersed setting. The primary and greatest thrilling submission of blockchain is electric money [16]. Persons rapidly understood the unchallengeable dispersed ledger knowledge, and dispersed notions behindhand blockchain can be extra modified and rummage-sale for numerous other requests like smart contracts, possessions label registries, digital voting, SCM, individuality management, Internet of Things, wireless sensor networks, digital possession managing, and numerous additional. Consequently, a substantial quantity of investigates and growth has since been ongoing on requests of blockchain, and additional new investigates are developing every day on conceivable upcoming applications [17].

4.1 Blockchain Used in WSN

Traditional wireless sensor-based network uses the centralized authentication method, and this method required trusted third party for the authorization process. Due to single point of authorization, these methods are under the threat of single point of failure [18]. Blockchain is the new way to provide the decentralized authentication process for the wireless sensor network. Table 3 shows the blockchain used in wireless sensor network to improve the energy efficiency of the network. We have summarized four core technical elements that enable the blockchain to provide trusted and secure services:

The first point is the distributed ledger that contains the particulars of the transactions dedicated on the blockchain. The transaction facts include speech of the receiver of the transaction, moved quantity, the timestamp, smart contract code, and its implementation effect. The transaction catalog is finished by construction manifold nodes at dissimilar places. Each node in the blockchain preserves pathway of the comprehensive ledger that does not permit the ledger information to get interfered. All the other nodes are official to screen the authority of the transactions [19]. The second is asymmetric encryption and authorization mechanism. All the transaction evidence stored on the blockchain is public. To safeguard the information security and individual confidentiality, the explanation individuality evidence is kept encoded and can be retrieved after the providing expert of the data proprietor [20]. The third part, called consensus mechanism, in which the effectiveness of the blockchain transaction in terms of stopping meddling is single minded by how all secretarial nodes spread consensus. The common consensus algorithms are proof of work (PoW), proof of stake (PoS), proof of authority (PoA), delegated proof of stake (DPoS), and proof of capacity (PoC) [21]. The last procedural division is the smart contract, which has roots in reliable and non-tampered data and is accomplished of automatically execution predefined codes by a blockchain miner. The

ledger standing on the blockchain system gets efficient by the results executed by smart contract. These fluctuations in the ledger cannot be interfered once they are established with a fixed consensus instrument as the gratified has been combined on the blockchain network [22]. The routing material can be learned by directing node which is not imperfect to its national routing node, and this can be done in an open, reliable, and distributed blockchain network. The proper usage of this routing information can improve the efficiency of the routing. The dynamic network is introduced with the concept of strengthening knowledge by some direction-finding structures. Reinforcement learning is a kind of machine learning procedure and characterizes as Q-learning which stretches response into the assortment of each step with prize and sentence apparatus [23]. The perception of reinforcement algorithm consists of five parts: environment, agent, state, action, and reward. An agent is capable of networking with the environment with the help of movements achieved. The condition of an agent is specified by the state constraint. Each state contains of numerous sets of movements that a manager can indicate. The agent can indicate one state achievement at a time and gets the reward results in the feedback on the achievement or failure of the act.

4.2 Blockchain in Agriculture

The use of data and statistics tries available to be increasingly crucial for the agroindustry zone to recover effectiveness and supportability. Walter in 2017 says information and communication technology (ICT) meaningfully enlarges the competence and efficiency of meeting, tapping gone, dismembering and employing material in cultivation. As per paper [22], horticultural authorities and humanizing systems to smoothly obtain update-to-date data and hence relax on healthier selections in their ordinary humanizing. For instance, Brown and Yousefi and Razdari, 2015, detected information on soil circumstances container provision agriculturalists' harvest organization cell phones reduce statistics price and therefore raise agriculturalists arrival to commercial subdivisions and economic assistance, and the progression of global positioning system (GPS) inspires logged preparation and computer hardware course and harvest traveling. For example, usual observant material is by and great supervision by united administration rudiments that consume their own benefit. They can operate the executive connected to information. The blockchain novelty helps to stock material and information that dissimilar performers and associates make all finished the entire value comprised process, after kernel to contract, of bringing an agricultural article [24]. Its assurances that the material and information are forthright to the comprised performers and associates, and each solitary documented data point is unchanging. A blockchain is a folder that covers timestamped bunches of connections and movements recognized with an article. Storage material in waiters centrally achieved by managers is additional at danger to damage and misrepresentation than allocating them to waiters on the Internet. The folder is unintelligibly valuable for making material ambitious moveable requests that assistance improves humanizing.

In calculation, the blockchain speeches the contest of creation an unfriendly accomplishment protected groundwork for Internet of Things and with various novelties exploited in ICT e-farming.

5 Conclusion

Blockchain is the technology used for the authentication and required to achieve digital currency. Blockchain has become very popular due to the extensive use of cryptocurrencies used to achieve digital currencies. Many of the applications are using blockchain to achieve streamlining processes, to achieve the efficiency, distributed security, data security, smart contracts, for the data sharing, health care, and manufacturing domains. In this paper, author established area and advantages of using blockchain in wireless sensor network and blockchain in agriculture. We have summarized four core technical elements that enable the blockchain to provide trusted and secure services. Still there are many limitations of blockchain in an open area including security, reliability, scalability, integration. Researchers have focus on these areas, but there are many practical and operative industrial applications.

References

1. Babich V, Hilary G (2018) Distributed ledgers and operations: what operations management researchers should know about blockchain technology. Manuf Serv Oper Manag (forthcoming) (in press). https://doi.org/10.2139/ssrn.3131250
2. Barnett BJ, Barrett CB, Skees JR (2008) Poverty traps and index-based risk transfer products. World Dev 36:1766–1785. https://doi.org/10.1016/j.worlddev.2007.10.016
3. Barnett BJ, Mahul O (2007) Weather index insurance for agriculture and rural areas in lower-income countries. Am J Agric Econ 89:1241–1247. https://doi.org/10.1111/j.1467-8276.2007.01091.x
4. Boysen N, de Koster R, Weidinger F (2019) Warehousing in the E-commerce era: a survey. Eur J Oper Res 277:396–411. https://doi.org/10.1016/j.ejor.2018.08.023
5. Brown ME (2015) Satellite remote sensing in agriculture and food security assessment. Procedia Environ Sci 29:307. https://doi.org/10.1016/j.proenv.2015.07.278
6. Caro MP, Ali MS, Vecchio M, Giaffreda R (2018) Blockchain-based traceability in agri-food supply chain management: a practical implementation. In: Proceedings of the 2018 IoT vertical and topical summit on agriculture-Tuscany (IOT Tuscany), Institute of Electrical and Electronics Engineers, Tuscany, pp 1–4
7. Chen W (2018) Administrative rules and regulations for the introduction of food information traceability in blockchain. Ph.D. thesis, Shanghai Normal University, Shanghai
8. Chod J, Trichakis N, Tsoukalas G, Aspegren H, Weber M (2019) On the financing benefits of supply chain transparency and blockchain adoption. Manag Sci (forthcoming) (in press). https://doi.org/10.2139/ssrn.3078945
9. Collier ZA, Bates ME, Wood MD, Igor L (2014) Stakeholder engagement in dredged material management decisions. Sci Total Environ 496:248–256. https://doi.org/10.1016/j.scitotenv.2014.07.044
10. Conradt S, Finger R, Spörri M (2015) Flexible weather index-based insurance design. Clim Risk Manag 10:106–117. https://doi.org/10.1016/j.crm.2015.06.003

11. Dalhaus T, Finger R (2016) Can gridded precipitation data and phenological observations reduce basis risk of weather index–based insurance? Weather Clim Soc 8:409–419. https://doi.org/10.1175/wcas-d-16-0020.1
12. Dalhaus T, Musshoff O, Finger R (2018) Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance. Sci Rep 8:46. https://doi.org/10.1038/s41598-017-18656-5
13. Finger R, Dalhaus T, Allendorf JS, Hirsch S (2018) Determinants of downside risk exposure of dairy farms. Eur Rev Agric Econ 45:641–674. https://doi.org/10.1093/erae/jby012
14. Gatteschi V, Lamberti F, Demartini C, Pranteda C, Santamaría V (2018) Blockchain and smart contracts for insurance: is the technology mature enough? Future Internet 10:20. https://doi.org/10.3390/fi10020020
15. Ge L, Brewster C, Spek J, Smeenk A, Top J, van Diepen F et al (2017) Blockchain for agriculture and food: findings from the pilot study. Wageningen Economic Research report; No. 2017-112. Wageningen Economic Research, Wageningen
16. Harz D, Gudgeon L, Gervais A, Knottenbelt WJ (2019) Balance: dynamic adjustment of cryptocurrency deposits. In: Proceedings of the 2019 ACM SIGSAC conference on computer and communications security. ACM, New York, NY, pp 1485–1502
17. Haveson S, Lau A, Wong V (2017) Protecting farmers in emerging markets with blockchain. Cornell Tech, Newyork, NY
18. Iansiti M, Lakhani KR (2017) The truth about blockchain. Harv Bus Rev 95:118–127. https://doi.org/10.3390/s19153267
19. IBM Institute for Business Value (2015) Device democracy: saving the future of the Internet of Things. Available at https://www.ibm.com/services/us/gbs/thoughtleadership/internetofthings/. Accessed 25 September 2019
20. Jha S, Andre B, Jha O (2018) ARBOL: smart contract weather risk protection for agriculture. Available at https://www.arbolmarket.com/wp-content/uploads/2018/09/ARBOL_WP-1.pdf. Accessed 15, 2018
21. Just RE, Calvin L, Quiggin J (1999) Adverse selection in crop insurance: actuarial and asymmetric information incentives. Am J Agric Econ 81:834–849. https://doi.org/10.2307/1244328
22. Kaddu S, Haumba EN (2016) Promoting ICT based agricultural knowledge management for increased production by smallholder rural farmers in Uganda: a case of Communication and Information Technology for Agriculture and Rural Development (CITARD), Butaleja. In: Proceedings of the 22nd standing conference of Eastern, Central and Southern Africa library and information associations (SCECSAL XXII), Butaleja, pp 243–252
23. Kamath R (2018) Food traceability on blockchain: Walmart's pork and mango pilots with IBM. J Br Blockchain Assoc 1:3712

EEG and ECG-based Drowsiness Detection: A Review on State of the Art

Asmita Sen Gupta, Monika Kumari, Shivanshi Shokeen, Anju Mishra, and Archana Singh

Abstract In this paper, we have collected, analyzed and presented a review of some of the latest research in the field of physiological signal-based drowsiness detection. The paper specifically focuses on presenting the state of the art in this direction by utilizing electroencephalogram (EEG) and electrocardiogram (ECG) signals. These methods are used for tracking and recording wave patterns of the brain and heart signals of an individual to find their drowsiness level which can further be used in developing many applications. One such application is in development of automated systems for identifying driver's drowsiness level while driving a vehicle. This work presents a thorough analysis method that has been implemented to find the drowsiness level of the person by recording the EEG and ECG signals of a person and what all solution have been proposed in the literature by different authors.

Keywords EEG · ECG · Drowsiness · Automated systems

1 Introduction

EEG and ECG are used for detection of drowsy behavior of the drivers. Also, it is used for those people who do repetitive tasks, that is, they work at both the day and night hours which are generally in occupations like traffic controllers and in the military environment. The most important use of this detection methodology is to prevent accidents to take place because of the drowsiness of the drivers which has taken place in recent years at a large scale. A person who is driving an automobile is a perplexing work which needs bodily and psychological consideration for awareness and give performance adequately and in most effective way. Work in shifts is present

A. S. Gupta (✉) · M. Kumari · S. Shokeen · A. Mishra · A. Singh
Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

A. Mishra
e-mail: amishra1@amity.edu

A. Singh
e-mail: asingh27@amity.edu

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_4

in many businesses (e.g., land transportation, manufacturing, military and maritime) to keep up 24 h activity. Broadened attentiveness for the duration of the evening late hours can create disruption between natural capacities and sociological needs, which thusly prompts drowsiness.

The term drowsiness normally alludes to a condition of laziness, tiredness and detachment that prompts the possibility that the person can get into a sleepy state. The essential arrangement is a state of mindfulness where an individual is absolutely prepared and can perform physical and mental endeavors while keeping up readiness and remaining on track. The ensuing class, NREM, is isolated into three stages; stage one is related to languor, while stages two and three are connected with light rest and significant rest (e.g., profound rest) states, separately. Tiredness is in the middle of state between awake and drowsiness. It is an advancement of psychophysiological state from alert toward rest causing issues in center, along these lines growing the reaction time. Consequently, this is the reason that laziness is unsafe in driving conditions, where a driver's deficiency of thought and extended reaction time can cause some serious driving incidents achieving veritable injuries and passing.

2 Literature Review

Detection of the Excessive Daytime Drowsiness in Resting State Electroencephalogram Recordings: Machine Learning Approach using EEG Sub-Bands and Channels by Breitenbach et al. [1] constructed a successful excessive daytime sleepiness (EDS) algorithm that basically divided many levels of the performance of people which depended on the experiment done of 8-minute unseen dormant state EEG with eyes in an open state. Random forest had 86.9% of accuracy. These outcomes were achievable via the blend of the sub-bands of EEG and a RF classifier. This model is a less-cost model and is quick in various medical field. Because of the subsequent low event of subjects with an ESS-score above the official threshold in the dataset, only patients with as light inclination to EDS could be compared with patients who are healthy which can be considered as a limitation.

EEG-based Internal Stress Observation: An overview of methods involved from EEG acquisition to system design by Jegan et al. [2] presents the planned investigation for building up an EEG sensor-based human inner stress perception framework. Here, we surveyed the strategies received for smart EEG signal obtaining from EEG sensor and have recorded the important features for identification of stress dependent on intellectual changes of brain action.

Novel Wearable EEG and ECG Measuring System for Stress Evaluation by Ahn et al. [3] was a study that presented a novel wearable gadget which can quantify EEG and electrocardiograms (ECG) all the while was intended to empower persistent stress levels in everyday life. The created framework is effectively worn by the two ears, is lightweight (i.e., 42.5 g) and displays an amazing noisy sound performance of 0.12 μVrms. Noteworthy time and recurrence highlights of EEG and HRV were

found in two distinct stressors, in particular the arithmetic mental tests and also Stroop color word, utilizing 14 subjects.

On-board Drowsiness Detection using EEG: Current Status and Future Prospects by Majumder et al. [4] study gave EEG examination for studies on drowsiness, current discoveries and future bearings of this domain which are quickly surveyed. Power spectral density (PSD)-based highlights are discovered to be the most normally utilized highlights for EEG-based studies on drowsy behaviors. EEG low- recurrence groups (alpha, delta and theta), particularly alpha band, exhibit an expansion in band power during the sleepy state contrasted with ready state. Conversely, high-recurrence groups (gamma and beta), explicitly beta band, exhibit a reduction in band power during tiredness. As far as cerebrum areas, occipital, parietal and frontal are intriguingly instructive, particularly, alpha from occipital and beta from frontal are two possible pointers groups which will assist with lessening the quantity of electrodes needed to build up a viable EEG-based recognition.

Effective mixture replica used for EEG-Based Drowsiness Detection by Budak et al. [5] study gives an effective strategy which proposed EEG-based laziness recognition. The proposed strategy utilizes three element extraction systems for hearty portrayal of the sleepiness EEG signals. In doing as such, power as well as nothing intersection distribution, ethereal entropy also, momentary recurrence highlights are separated in one of the first removal systems. The second component removal machine takes bottomless highlights as of pre-prepared AlexNet also, models VGG16. At last step, the measurable highlights of quick recurrence of the TQWT-decayed electroencephalogram signals be taken into the third element removal machine. The highlights are after that either arranged separately by LSTM network or the yields of every LSTM network be consolidated on a major vote system.

Drowsiness recognition utilizing pulse changeability examination dependent on microcontroller unit by Hendra et al. [6] shows studies for detecting drowsiness utilizing HRV examination dependent on microcontroller part. ECG signal is acquired via AD8232 unit and prepared in microcontroller part. ECG can be recorded for the duration of the topic utilizing dynamic test system. Remove highlights as of HRV and utilize neural network and radial basis to group between drowsy and alert state.

Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio by Mehta et al. [7] which presents a continuous framework that screens and identifies the deficiency of recognition by people driving is proposed. The drivers have been identified by catching face aspect ratios, and caution is given to the driver to maintain a strategic distance from continuous accidents.

A Review on EEG-Based Automatic Sleepiness Detection Systems for driver by Balandong et al. [8] study gave more secure conditions in land transportation, which has been thoroughly investigated in laboratories. Any detoriation in the process is removed by identification of the drowsiness. Discoveries show that mind action as evaluated from scalp EEG can be used for estimating drowsiness levels. Consistent improvement to locate the best blend between scalp zones, EEG-based highlights and classifiers is utilized to improve electroencephalography-based sleepiness detection system (ESDS) classification and application in the various fields.

Abnormal EEG signals automated identification through a Deep Convolutional Neural Network Model by Yıldırım et al. [9] is an examination, another profound convolutional neural network, that is, one-dimensional in nature and is also referred to as 1D-CNN model is recommended hugely for the measurement of EEG signals that is specific as well as unusual in nature. The recommended model which without the utilization of any component extraction orders the EEG signals and also has an end-to-end arrangement. In this experiment, the EEG signals from transient to occipital (T5-O1) are single-channel (TUH) EEG abnormal corpus (v2.0.0), that is, EEG dataset from Temple University Hospital to build up 1D-CNN prototype. The created model has resulted in 20.66% rate of errors in grouping the typical and unusual EEG signals.

ECG and EEG Signals used for Correlation Analysis between Cardiac and Brain Activity by Kerwin et al. [10] presents an examination inspects of the correlation between cardiovascular capacity and mind action during occasions of physical furthermore, mental pressure. Eight subjects took an interest in the lead of the analysis comprising of 4 girls and 4 boys and were given assignments, for example, sitting without any movement, running along with watching exciting video cuts, while ECG and EEG sensors were worn by them. The outcomes indicated an expansion in pulse as well as mind action (alpha waves, beta waves, delta waves, as well as theta waves) upon boost initiation. The investigation recommends that some cerebrum action that is related with a specific stimulus is revived by various different types of operations.

Hybrid approach for detecting Drowsiness of drivers using physical signals for improvement in System execution and Wearability by Awais et al. [11] study gave utilization of feature extraction used for recognition of basically languor of drivers which actually had been implemented in 22 subjects where a test system-based driving condition was created. The ECG recordings were measured which includes time space complexity, power which is relative characteristics were taken and were approved by video data. Moreover, EEG frequency-domain investigation indicated basic languor-related changes in absolute power in the delta, theta and alpha groups in the central, parietal and occipital zones. Power which is relative exhibited immense changes in the alpha band in occipital and parietal zones. The results revealed an immense expansion in both supreme and relative forces during the transitory stage from the alert state to the sleepy state (Fig. 1).

Detecting drowsy behavior using heartbeat rate variableness by Laguna et al. [12] is a study that showed 9.9–29.9% street casualties identified with driving in a sleep state. The sluggishness location dependent mostly on the natural and signals from the various automobiles which are concentrated for the vehicle wellbeing. Self-governing sensory system movement, mostly, is estimated without putting tools into the body taken from pulse inconstancy and features got by using ECG gives adjustments mostly in tension, outrageous weakness, sluggishness scenes. They guessed that these adjustments show on HRV, and in this way, it is utilized for distinguishing languor of drivers.

Fig. 1 Simple EEG process for drowsiness detection where from signals, data is extracted by various algorithms and then it is deduced whether the person is in drowsy state or awake state

Detection of drowsy driving by Analysis Using Electroencephalogram, K-Means Clustering, Wavelet transformation by Gurudath et al. [13] builds programmed framework for driving in languor state identification and investigating brain signs. Transformation in the wavelet a powerful device used in giving an analysis of the time just as frequency parts covered up in such non-fixed signs. In genuine driving situations where calculation of EEG signal manually is unfathomable what is more, absurd, the K-means grouping method is a solid gadget to conclude which is laziness-related state: (i.e., alert, tired or resting) every model has a spot.

It is assumed that classes procured by bunching the data difference give the best of gatherings when obviously evaluated.

Electroencephalogram based Detection of drowsiness for driving safely using Chaotic characteristics and Tests based on statistics by Mardi et al. [14] study gave a simple procedure for accumulation of EEG information from drowsy drivers. A recording of EEG information was done in virtual driving circumstance, a domain that is made dependent on the real world. This virtual driving circumstance was a laboratory circumference. Members were not allowed to sleep for 19.9 h before the actual examination was to be conducted. Subsequent to measurement of signals and furthermore imaging signals, simultaneously, preparing were done on databases, and extracting the features was done for the subsequent stage. The accuracy is above 83% in this process. This new convention would be an appropriate and straight easy path for information collection in virtual driving condition.

EEG Signals classification regarding Detection of drowsiness in Computer Interface and brain by Kumar et al. [15] is study about that drowsiness detection system depends on distinguishing reasonable to related to driving potentially related to automobile factors which corresponded to degree of the languor. Mind PC framework empowers management of gadgets and correspondence w.r.t different people, just by brain movement, and not utilizing muscles. Primary implementation regarding brain computer interface is an innovation mainly involving debilitated people. Models encompassing gadgets which are constrained using brain computer interfaces are counterfeit appendages, gadgets for spelling, condition management frameworks. Framework utilizes EEG.

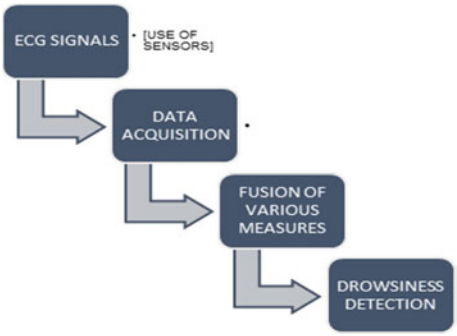


Fig. 2 Simple ECG process for drowsiness detection which from signals, data is extracted by various algorithms and then it is deduced whether the person is in drowsy state or awake state

Electrocardiogram Sensors for detecting Drowsiness of drivers by Gromer et al. [16] gives a solution of Android for the calculation for driver's sleepiness. Progress depends on a custom PCB plan that catches the ECG. PCB is to be connected for a module to an Arduino used in quick analysis. Accordingly, it is conceivable to say something about exhaustion of the driver. On account of a sensor combination, this can add to a quantified articulation of the driver's condition (Fig. 2).

3 Methodology

- Gathered various research papers on our topic and analyzed them.
- Found out various algorithms like K-means clustering, CNN algorithm and random forest classifiers.
- Analyzed them by defining some parameters.
- Parameters that were used were no. of electrodes used, dataset used, no. of participants and the recording time.
- We took the research papers from the year 2020, 2019, 2018, 2017, 2016, 2014 and 2011.
- On the basis of our analysis, we deduced which methodologies are better.
- The summary and analysis are presented in Table 1. From the table, we can summarize that many methods have introduced in for detection of drowsiness.

4 Conclusion

This research paper has presented an analytical review of various research papers. Our main objective of this paper was to find the best technology which would be useful for drowsiness detection by taking into consideration various parameters like no. of participants, datasets, recording time, no. of electrodes used. We have also studied about EEG and ECG. We have seen in our research that the no. of participants in the research on abnormal EEG signals automated identification through a deep

Table 1 Summary of latest research work for drowsiness detection from EEG/ECG signals

S. No	Paper title	Year	Author	No. of electrodes	Datasets	No. of participants	Recording rate
1	Detection of the Excessive Daytime Drowsiness in Resting State Electroencephalogram Recordings: Machine Learning Approach using EEG Sub-Bands and Channels	2020	H. Baumgartl, J. Breitenbach, R. Buettner	62	Lepzig Mind-Brain-Body database	210	8 min
2	EEG Based Internal Stress Observation: An overview of methods involved from EEG acquisition to system design	2020	P. Subha Hency Jose, P. Rajalakshmy, V. H. Ashwin, R. Jegan	10–20	EEG during mental arithmetic tasks database (PhysioNet)	2–5	5–10 min
3	A Novel Wearable EEG and ECG Measuring System for Stress Evaluation	2019	Kim Hee Chan, Yunseo Ku and Ahn Joong Woo	4	HRV	14	80 s
4	On-board Drowsiness Detection using EEG: Current Status and Future Prospects	2019	N. Chunwu Wang, B. Guragain, S. Majumder	21	EEG datasets	5–10	5–10 min
5	Effective mixture replica used for EEG-Based Drowsiness Detection	2019	V. Bajaj, U. Budak, Y. Akbulut, O. Atilla and A. Sengur	0	MIT-BIH Polysomnographic database	2–5	5–10 min

(continued)

Table 1 (continued)

S. No	Paper title	Year	Author	No. of electrodes	Datasets	No. of participants	Recording rate
6	Drowsiness recognition utilizing pulse changeability examination dependent on microcontroller unit	2019	D. Kurniawan, M. Hendra, C. Trio, Riza Vina P. Utomo, Nuryani Nuryani	3	HRV	4	10 min
7	Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio	2019	S. Dadhich, S. Mehta, S. Gumber, A. Jadhav Bhatt	0	Eye aspect ratio and eye closure ratio datasets	2	5–10 min
8	A Review on EEG-Based Automatic Sleepiness Detection Systems for driver	2018	R. Fayyaz Ahmad, R. Petrus Balandong, A. Saeed Malik, M. Naufal Mohamad Saad	2–30	EEG datasets	5–10	5–10 min
9	Abnormal EEG signals automated identification through a Deep Convolutional Neural Network Model	2018	U Rajendra Acharya, Baran Baloglu, Ö. Yıldırım	19	TUH EEG dataset	2717	60 s

(continued)

Table 1 (continued)

S. No	Paper title	Year	Author	No. of electrodes	Datasets	No. of participants	Recording rate
10	ECG and EEG Signals used for Correlation Analysis between Cardiac and Brain Activity	2018	Rhen Anjerome R. Bedruz, Shaina Mara D. Caguicla, K. Martin S. Ilagan, K. Richelle C. Monsale, A. Gerard G. Santos, R. Kerwin C. Billones Ira C. Valenzuela, Jayvee P. Villanueva, Elmer P Dadios	3	EEG and ECG datasets	18	2 min
11	Hybrid approach for detecting Drowsiness of drivers using physical signals for improvement in System execution and Wearability	2017	M. Awais, M. Drieberg and N. Badruddin	5	HRV, fatigue sheet, Stanford sleepiness scale, eye blinking dataset, head movement dataset	22	80 min
12	Detecting drowsy behavior using heartbeat rate variableness	2016	P. Laguna, R. Bailon, A. Bartra, J. Vicente	–	Three driving databases which were lack of slumber, which was clarified as languor or conscious	30	3475 min
13	Detection of drowsy driving by Analysis Using Electroencephalogram, K-Means Clustering, Wavelet transformation	2014	H. Bryan Riley, N. Gurudath	22	PhysioNet sleep-EDF database	12	5–10 min

(continued)

Table 1 (continued)

S. No	Paper title	Year	Author	No. of electrodes	Datasets	No. of participants	Recording rate
14	Electroencephalogram based Detection of drowsiness for driving safely using Chaotic characteristics and Tests based on statistics	2011	S. Naghmeh, Z. Mardi, M. Ashtiani, and M. Mikaili	22	Higuchi's fractal dimension, Petrosian's fractal dimension datasets	10	45 min
15	EEG Signals classification regarding Detection of drowsiness in Computer Interface and brain	2011	Dr. S. V Prasada Raju, G. R. Kumar, D. Santhosh Kumar	3	Traffic safety records, blink rate, BCI datasets	5–10	5–10 min
16	Electrocardiogram Sensors for detecting Drowsiness of drivers	2011	R. Seepold M. Gromer, D. Salb, T. Walzer, NM Madrid	3	HRV, QRS	5–10	4–5 min

convolutional neural network model, i.e., 2717. Also, the recording time for each participant is just 60 s. So, we can see in here that as participants are high, so the accuracy is also gradually increasing. The created model has resulted in 20.66% rate of errors in grouping the typical and unusual EEG signals which are quite less. Also, the TUH EEG dataset used in this research can also be used for further studies regarding this topic. Also, 1D-CNN is one of the best methods used for anomaly, motor-error detection and many more. Also, it is less expensive as compared to other techniques as it has easy and strong configuration of 1D. Other methods were also quite effective and took less time, but this method after our research has been concluded to be most effective. By our study, we can say which method is the most useful ones and can make more development on the pre-existing procedures.

5 Future Scope

The use of EEG and ECG for drowsiness detection has always played a vital role. So far, EEG and ECG have helped a lot when we wish to check the drowsiness level of the individual. The only change required is the use of no. of electrodes. Use of electrodes for EEG and ECG detection is one of the most important techniques. The amount of the electrodes used should be less. Also, a better framework for placing the electrodes should also be kept in mind. This framework may be a cap or a helmet. The accuracy levels for all the methodologies have also been studied. Work should be done toward increasing the accuracy levels of the entire process. Working with new algorithms is also needed to be explored to check whether these algorithms may take the accuracy level at a much higher rate. Also, new features should be added to provide more assistance to people. These features may include chatbot which may help people to talk to the doctors when they require some assistance regarding their health issues. If these requirements are met in the future, a lot more help can be provided to the health care industry.

References

1. Breitenbach J, Baumgartl H, Buettner R (2020) Detection of excessive daytime sleepiness in resting-state EEG recordings: a novel machine learning approach using EEG sub-bands and channels. In: Americas conference on information systems
2. Jegan R, Ashwin VH, Rajalakshmy P, Subha P, Jose H (2019) EEG based internal stress observation: an overview of methods involved from EEG acquisition to system design. J Crit Rev
3. Ahn JW, Ku Y, Kim HC (2019) A novel wearable EEG and ECG recording system for stress assessment. In: National center for biotechnology information. U.S. National Library of Medicine
4. Majumder S, Guragain B, Wang C, Nicholas (2019) On-board drowsiness detection using EEG: current status and future prospects. In: IEEE international conference on electro information technology (EIT)

5. Budak U, Bajaj V, Akbulut Y, Atilla O, Sengur A (2019) An effective hybrid model for EEG-based drowsiness detection. IEEE Sens J
6. Hendra M, Kurniawan D, Chrismiantari RV, Trio, Utomo P, Nuryani N (2019) Drowsiness detection using heart rate variability analysis based on microcontroller unit. In: 9th international conference on physics and its applications (ICOPIA)
7. Mehta S, Dadhich S, Gumber S, Bhatt AJ (2019) Real-time driver drowsiness detection system using eye aspect ratio and eye closure ratio. In: International conference on sustainable computing in science, technology and management (SUSCOM)
8. Balandong RP, Ahmad RF, Saad MNM, Malik AS (2018) A review on EEG-based automatic sleepiness detection systems for driver. IEEE Access
9. Yıldırım Ö, Baloglu UB, Acharya UR (2018) A deep convolutional neural network model for automated identification of abnormal EEG signals. In: Springer on neural computing and applications
10. Billones RKC, Bedruz RAR, Caguicla SMD, Ilagan KMS, Monsale KRC, Santos AGG, Valenzuela IC, Villanueva JP, Dadios EP (2018) Cardiac and brain activity correlation analysis using electrocardiogram and electroencephalogram signals. In: IEEE 10th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM)
11. Awais M, Badruddin N, Drieberg M (2017) A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. In: National library of medicine
12. Vicente J, Laguna P, Bartra A, Bailon R (2016) Drowsiness detection using heart rate variability. In: Springer on medical and biological engineering and computing
13. Gurudath N, Riley HB (2014) Drowsy driving detection by EEG analysis using wavelet transform and K-means clustering In: The 11th international conference on mobile systems and pervasive computing. School of Electrical Engineering and Computer Science, Ohio University, Athens, Ohio 45701, USA
14. Mardi Z, Ashtiani SNM, Mikaili M (2011) EEG-based drowsiness detection for safe driving using chaotic features and statistical tests. In: National center for biotechnology information, U.S. National Library of Medicine
15. Kumar GR, Raju SVP, Kumar DS (2012) Classification of EEG signals for drowsiness detection in brain and computer interface. GESJ Comput Sci Telecommun 4(36)
16. Gromer M, Salb D, Walzer T, Madrid NM, Seepold R (2011) ECG sensor for detection of a driver's drowsiness. In: IEEE conference

Attendance Management System Using Face Recognition

Vivek Mishra, Rashmi Mishra, Rupa Rani, and Irfan Khan

Abstract Face recognition has been a hot topic of research for a long time. Face recognition is a difficult problem to solve and that is what attracts researchers toward it. Automated attendance system is a practical application of face recognition technology. As we know, we can identify any person by just watching his face, so we are trying to develop automatic attendance system based on face recognition. In this paper, we have used high-definition camera, and video captured by camera is being used by our model to detect and recognize human faces. With this approach, we are able to recognize students in the class, and their attendance is being marked in a database. In this paper, we are providing a user interface which will help the faculty to mark the student's attendance from their computer screen.

Keywords OpenCV · LBPH · Face recognition · Eigenfaces · Fisherfaces

1 Introduction

In this modern era, where everything is on the way of being digitalized attendance in schools/colleges is still recorded manually which makes this task very outdated, time-consuming, and tedious. Manual way of marking attendance also faces problems like proxy and wastage of paper. With the advancement of technology, different methods to replace the manual attendance system have been used. These methods include RFID-based system, IRIS recognition-based system, and fingerprints-based systems.

V. Mishra
ABES Engineering College, Ghaziabad, Uttar Pradesh, India

R. Mishra (✉) · R. Rani · I. Khan
Krishna Engineering College, Mohan Nagar, Ghaziabad, Uttar Pradesh, India

R. Rani
e-mail: rupa.rani@lloydcollege.in

I. Khan
e-mail: irfan.khan@lloydcollege.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_5

These systems have their own drawbacks as they are based on queue which is time-consuming. IRIS and fingerprints-based system are considered as privacy threats. In our paper, we have used face recognition approach for our attendance system. There are various approaches for face recognition such as neural network, artificial neural network, CNN. These are some of deep learning methods. In our paper, we are using OpenCV-based approach which is based on machine learning.

Face recognition is a difficult task due to varying lightning condition, movement, and pose variation. In this, we are using LBPH algorithm for face recognition which gives better results in different lightning condition.

2 Literature Survey

Objective of our paper is to develop an automated attendance system which will replace the traditional way of marking attendance in classroom. Automated attendance system is a practical application of face recognition technology which solves the real-world problem.

Face detection and recognition are two main components of automated attendance system. There are basically two techniques for face detections: feature-based and image-based. Feature-based approach uses different types of features such as skin color-based features as suggested in [1], grayscale-based features described in [2], and edge-based features described in [3]. Skin color is an important feature of human face. Colors are easy to process as compared to other facial features. Gray information of a face such as eyebrows and lips which generally appears darker is important features. Many algorithms use these features. Sakai et al. 1972 introduced face detection based on edges. Basically, it was based on analyzing line drawings on human faces from images to locate facial features. Image-based approach uses mainly neural network and statistical methods.

In this system, we have used Haar cascade classifier for face detection which is based on Viola-Jones [4] framework which was proposed by *Paul Viola and Michael Jones* in 2001 and is available in OpenCV. LBPH algorithm is being used for face recognition. LBPH algorithm was first described in 1994 since then it has been used extensively.

Kawaguchi in 2005 [5] proposed a lecture attendance system. They used continuous monitoring method for face recognition. In this method, they used two cameras, one for taking images and other for recording the position of student in the class. In 2012, another paper [6] proposed student attendance system based on face recognition. They used PCA for face recognition and maintained a log for student in time and out time. In [7], author has proposed a smart attendance system based on facial recognition. This method uses LBPH algorithm for face recognition which is provided by OpenCV library. In [8] this paper, author has used face recognition in real-time background world for their attendance monitoring system. They have used PCA algorithm. The main problem with this technique is background subtraction from image which is still a challenge. In [9], author performed comparison among

KNN, CNN, and combination of LBP and CNN. Combination of LBP and CNN gives out 90% accuracy, while KNN and CNN give 77.27% and 80.303% accuracies, respectively. In [10], author has suggested automated attendance using Haar classifiers, KNN, CNN, SVM, generative adversarial networks, and Gabor filters. In [11], author has proposed automated attendance system using LBPH algorithm for face recognition and has recorded 77% accuracy in live video. We have also studied about the existing attendance system which is as follows:

2.1 RFID-Based System

In the research paper [12], author has suggested RFID system in which student has to show his RFID card in front of RFID reader to mark his presence. The drawbacks of this system are that it is time-consuming and also it creates disturbance in the class during attendance.

2.2 IRIS Scan-Based System

In [13], author has described IRIS scan-based system in which the uniqueness of IRIS is being utilized for creating attendance system. It is quite efficient method. But, the drawback of this system is that IRIS scanning-based system is considered as privacy invasion, so many people tend to avoid it.

2.3 Fingerprints-Based System

In [14], author has described fingerprints-based system. Each individual has different fingerprints, and attendance systems based on fingerprints are already in use in various schools/colleges. It also has same drawback as IRIS-based system, it also violates privacy of individual. It is also time-consuming, and it does not support as contactless system. Table 1 shows the comparison of existing systems.

Table 1 Comparison of different approaches

Feature	RFID-based	IRIS scan	Finger prints
Privacy violation	No	Yes	Yes
Contactless	No	Yes	No
Fraudulent access	Possible	No	No
Queueing	Yes	Yes	Yes

Table 2 Comparison of OpenCV algorithms

Criteria	Eigenfaces	Fisherfaces	LBPH
1. Principle of dataset generation	Component-based	Component-based	Pixel-based
2. Basic principle	PCA	LDA	Histogram
3. Background noise	Maximum	Medium	Low
4. Sensitivity to light	High	High	Low

We have also studied about the algorithms available in OpenCV for face recognition. There are three algorithms available in OpenCV which are Fisherfaces, eigenfaces, and LBPH algorithm. Fisherfaces use linear discriminant analysis (LDA) technique, eigenfaces use principal component analysis (PCA), while LBPH algorithm uses histogram as basic principle. Table 2 shows [15] study of these three algorithms.

So, conclusion of our literature survey is that existing attendance systems have some drawbacks associated with them, so we have proposed an automated attendance system which has following merits over current system: Proposed system is contactless, time-saving and secure, and no disturbance during attendance in the class.

3 Proposed Methodology

Figure 1 shows that our automated attendance system works in the following steps.

3.1 Image Acquisition

During registration of student, he will have to face a camera dedicated for registration purpose only. These images will be stored along with student's details. Around 50–60 images will be taken per student for training the model. During attendance, images will be taken from video recorded by cameras installed in classes. We will record video of short duration 2–3 times so that no student will be missed from face recognition.

3.2 Face Detection

Faces will be detected from the recorded videos. Our system uses Viola-Jones algorithm for face detection. This algorithm adapts Haar-like features for face detection. Figure 2 shows Haar-like features. AdaBoost algorithms help this algorithm to become a strong classifier.

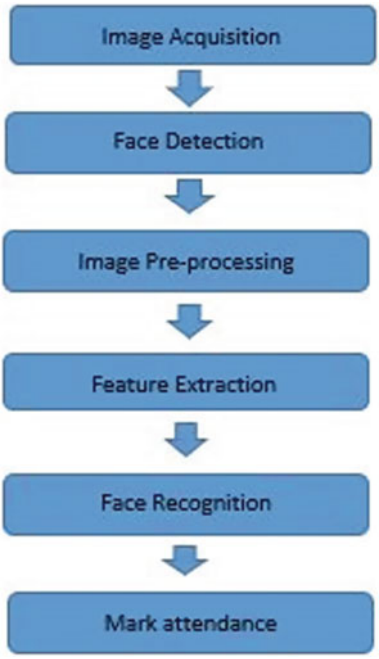


Fig. 1 Steps in implementation

Fig. 2 Haar-like features

3.3 Image Preprocessing

In this step, images are resized according to the requirement. Images from RGB format will be converted to grayscale. Grayscale images are easy to process and extract features. Figure 3 shows results of this step.

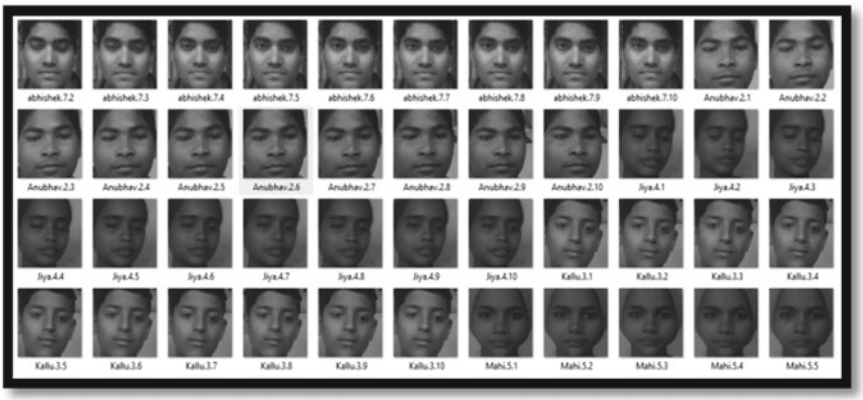


Fig. 3 Cropped face images in grayscale format

3.4 Feature Extraction

LBPH algorithm uses LBP operator for feature extraction. LBP operator is a simple but very powerful texture operator. LBP operator has following parameters as discussed in [16]:

Radius: It is the radius around central pixel, and it creates circular local binary pattern.

Neighbors: It represents the number of sample points to build the circular binary pattern, and more the number of neighbors, more will be computational cost.

Grid X: It represents number of cells in horizontal direction. It is generally set to 8. Higher number of cells gives finer grid with increased dimensions of feature vector.

Grid Y: It represents number of cells in vertical direction. It is generally set to 8. Higher number of cells gives finer grid with increased dimensions of feature vector.

LBP Operation: [17] Eq. (1) gives the formal description of LBP operator where (X_C, Y_C) is central pixel intensity. I_P and I_C are being intensity of neighbor pixels.

$$\mathrm{LBP}(X_c, Y_c) = \sum_{p=0}^{P-1} 2^p s\left(i_p - i_c\right) \tag{1}$$

S is the sign function given by Eq. (2). It is used for creating binary matrix with the help of central value.

$$S(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

For a given point (X_C, Y_C), the position of $(X_{P,} Y_P)$, $p \in P$, neighbor is calculated by Eqs. (3) and (4).

$$X_p = X_c + R\text{Cos}\left(\frac{2\pi p}{P}\right) \tag{3}$$

$$Y_p = Y_c - R\text{Cos}\left(\frac{2\pi p}{P}\right) \tag{4}$$

where R is the radius of circle and P is sample points. In [16], author shows LBP operation. In LBP operation, the original image is converted into a transitional image which will outline the facial region. Figure 6 shows the working steps of LBP algorithm. LBP operator creates 3×3 window which is described by Fig. 4. Each cell of window represents pixel intensity. We can represent window as 3×3 matrix. Each cell represents intensity of each pixel in the range of [0, 255]. A binary matrix is created by comparing the central value with all neighbors. If the value of neighbor is greater than central value, then it is converted to 1 else 0. After this binary value is concatenated to give 8-bit binary number, this binary number is then converted to decimal representation which gives the central value of matrix. Figure 4 shows this process [18]. After appending the binary numbers line by line, we get (10,001,101) which is equivalent to 141. It represents the central value of matrix which is a pixel value of original image.

This process continues for every position. At the end, we get a new image which represents better characteristics. After this step, we extract histogram of each region. Now, we have image in grayscale, so values of histogram will range from 0 to 255 representing occurrence of every pixel intensity. Figure 5 shows [18] calculation of histogram. Now, all the histogram will be combined to give a new histogram which will represent the characteristic of original images.

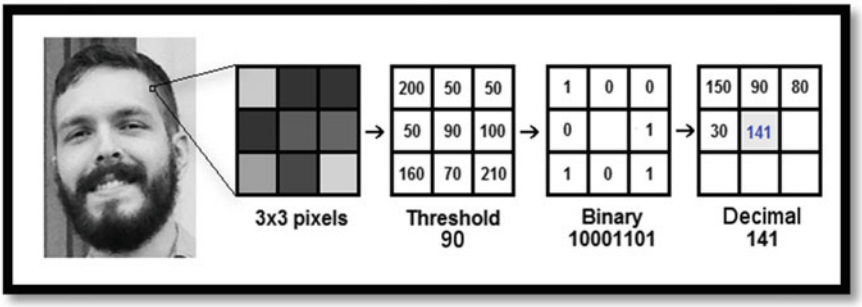


Fig. 4 LBP operation

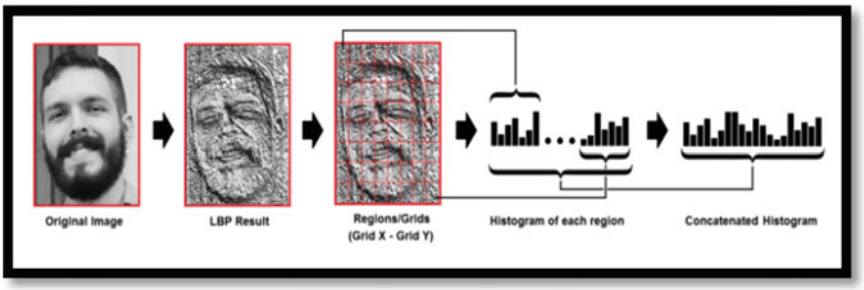
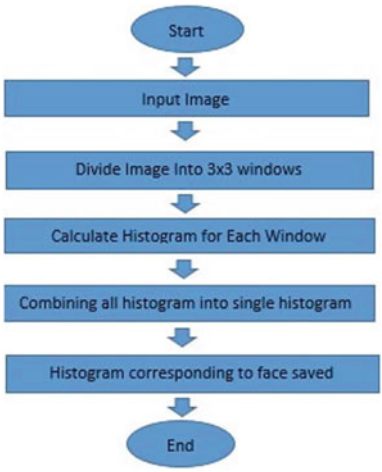


Fig. 5 Histogram calculation

Fig. 6 Flowchart of LBPH algorithm during training



Face Recognition: After training the algorithm, we are left with histograms corresponding to each face. When a test image is given to system, its histogram will be created in the same way as described above.

This histogram will be compared with all existing histograms, and the output from the algorithm will be the ID of closest matching histogram.

We can compare two histograms by using different method such as Euclidean distance, chi-square, absolute value. In this paper, we will be using Euclidean distance. Equation (5) represents Euclidean distance formula.

$$D = \sqrt{\sum_{i=0}^{i=n} (\text{hist}_{1i} - \text{hist}_{2i})^2} \tag{5}$$

Distance D is the distance between two histograms. It is also called confidence. Lower the D, closer are two histograms.

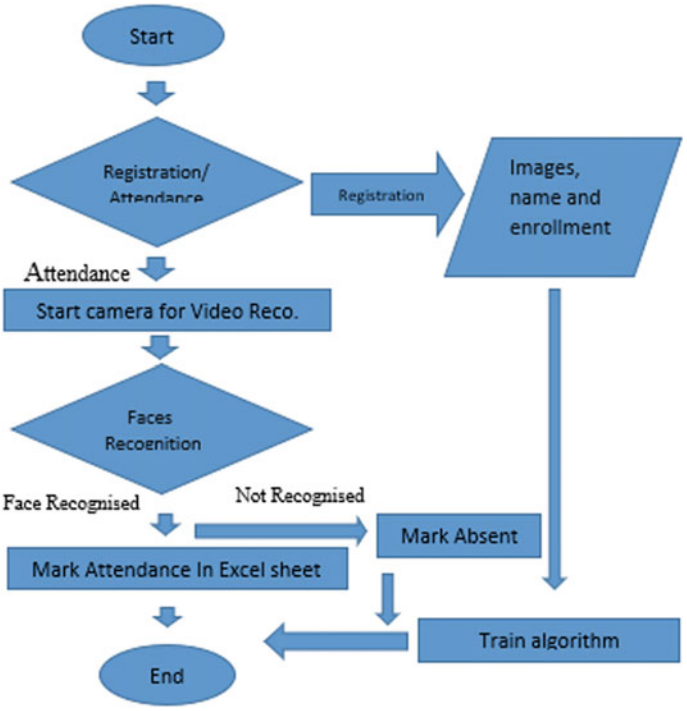


Fig. 7 Flowchart of model

Mark attendance: This is the final step in our system. In this step, recorded video will be given to our system, faces of students will be recognized, and they will be marked as present. Our application will enroll new student with his name, enrollment number, and images and then model will be trained. Figure 7 shows the flowchart of our model.

4 Results

We have implemented the above proposed approach for automated attendance system. We have created an application by using Python 3.6. Figure 8 shows the interface of our model. GUI has been created by using Tkinter library of Python. Our system detects face of student recognize it and then marks his attendance in an Excel sheet. The application has different modules for capturing images, training model, and for taking attendance. Figures 9 and 10 are showing how attendance is being marked of an individual person and group of persons, respectively. Figure 11 is showing the results. Figure 12 shows the confirmation of attendance by GUI pop up. Table 3 shows the results of our experiment. While using static images for

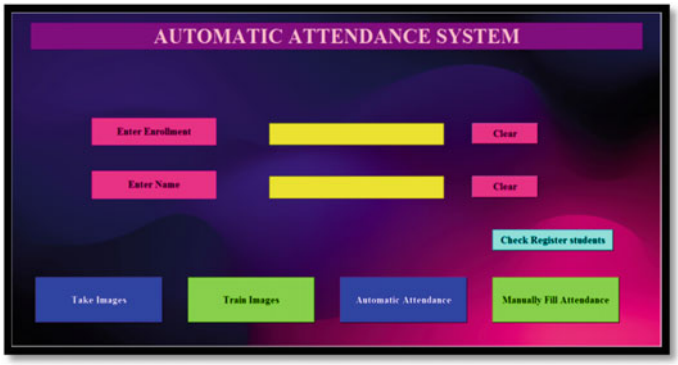


Fig. 8 Interface for face recognition-based attendance system

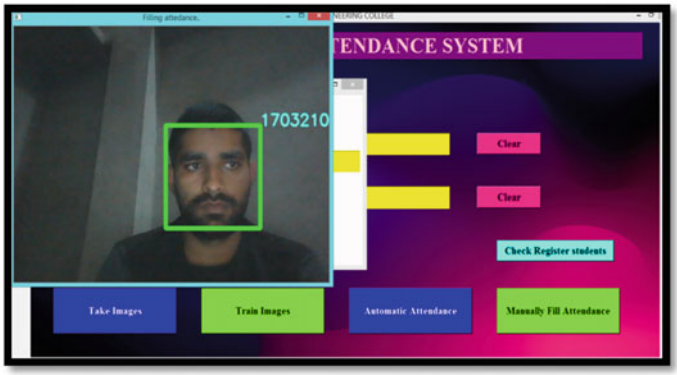


Fig. 9 Marking attendance of single person

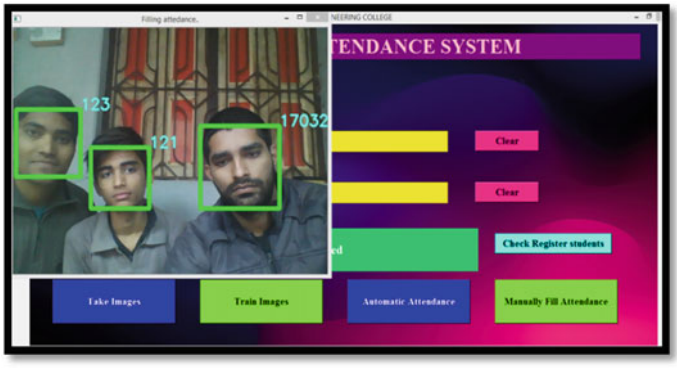


Fig. 10 Group attendance

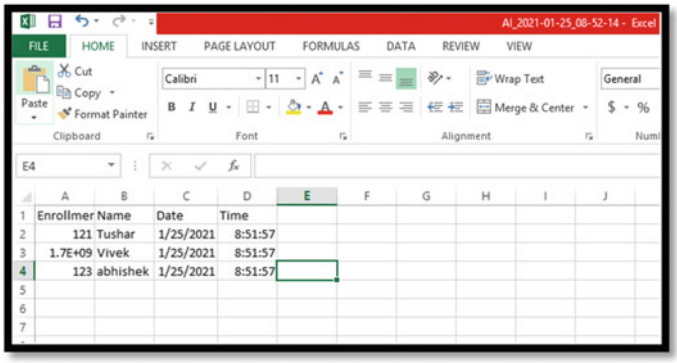


Fig. 11 Group attendance marked in Excel

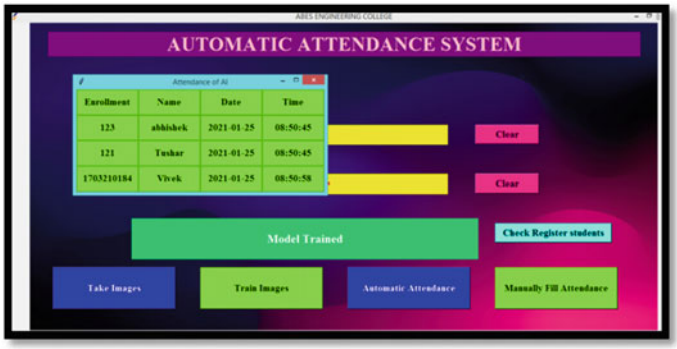


Fig. 12 GUI showing student's detail whose attendance has been marked

Table 3 Accuracy comparison of LBPH, eigenfaces, and Fisherfaces

Dataset	LBPH (%)	Eigenface (%)	Fisherface (%)
Recognition accuracy (static images)	93.33	91.66	90
Recognition accuracy (live camera)	79.16	62.5	58.33

testing, accuracy of our model using LBPH algorithm was 93.33%; on the other hand, accuracy decreases to 79.16% when we performed this experiment in uncontrolled environment using live camera. Similarly, we have performed experiments with eigenface and Fisherface algorithms, and results are given in Table 3.

5 Conclusion and Future Scope

From our experiment, we found out that LBPH algorithm outperforms other two algorithms in terms of face recognition rate. LBPH algorithm is more robust than other two and performs better in in different light condition. Real-time face detection is very difficult task to perform due to varying background, moving faces, and changes in light condition. With advancement of the face recognition technology, automated attendance system is a solution of real-world problem. Our paper shows that it is possible to record attendance by using face recognition method. It can be a very useful tool for any school/college or any institute which has manual system of attendance. It gives us a secure and time-saving tool. This application gives use an idea how automated system will work. Efficiency of the system can be increased in the future with more efficient algorithms. This application can be integrated with a database and can be used to generate reports of student's attendance.

Acknowledgements It gives us immense pleasure to acknowledge that this paper was completed under the guidance of Ms. ***Rashmi Mishra*** (Assistant Professor). She guided us with her knowledge every time we faced any challenge.

References

1. Alabbasi HA, Moldoveanu F (2014)Human face detection from images, based on skin color. In: 2014 18th international conference on system theory, control and computing (ICSTCC), Sinaia, Romania, pp 532–537. https://doi.org/10.1109/ICSTCC.2014.6982471
2. Kurylyak Y, Paliy I, Sachenko A, Chohra A, Madani K (2009) Face detection on grayscale and color images using combined cascade of classifiers. Int J Comput 8:61–71
3. Sikarwar R, Agrawal A, Kushwah RS (2015) An edge based efficient method of face detection and feature extraction. In: 2015 fifth international conference on communication systems and network technologies, Gwalior, India, pp 1147–1151. https://doi.org/10.1109/CSNT.2015.167
4. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, Kauai, HI, USA, pp I–I
5. Kawaguchi Y (2005) Face recognition based lecture attendance system. In: 3rd AEARU, October 2005
6. Kar N, Debbarma MK, Saha A, Pal DR (2012) Study of implementing automated attendance system using face recognition technique. Int J Comput Commun Eng 1(2):100–103
7. Bussa S, Mani A, Bharuka S, Kaushik S (2020) Smart attendance system using OPENCV based on facial recognition. Int J Eng Res V9. https://doi.org/10.17577/IJERTV9IS030122
8. Sayeed S, Hossen J, Kalaiarasi SMA, Jayakumar V, Yusof I, Samraj A (2017) Real-time face recognition for attendance monitoring system. J Theor Appl Inform Technol 95(1)
9. Sawardekar S, Naik SR (2018) Facial expression recognition using efficient LBP and CNN. Int Res J Eng Technol (IRJET) 5(6). e-ISSN 2395-0056 p-ISSN 2395-0072
10. Dev S, Patnaik T (2020) Student attendance system using face recognition. In: 2020 international conference on smart electronics and communication (ICOSEC), pp 90–96.https://doi.org/10.1109/ICOSEC49089.2020.9215441
11. Chinimilli BT, Anjali T, Kotturi A, Kaipu VR, Mandapati JV(2020) Face recognition based attendance system using Haar cascade and local binary pattern histogram algorithm. In: 2020

4th international conference on trends in electronics and informatics (ICOEI) (48184), pp 701–704. https://doi.org/10.1109/ICOEI48184.2020.9143046
12. Mistri RK (2018) RFID based attendance system. Int J Res Appl Sci Eng Technol 6:2237–2241. https://doi.org/10.22214/ijraset.2018.5364
13. kokpujie O, Kennedy O et al (2017) Design and implementation of a student attendance system using iris biometric recognition. In: 2017 international conference on computational science and computational intelligence (CSCI). IEEE
14. Singh S, Mishra P, Yadav D (2015) Fingerprint based attendance system using microcontroller and labview. Int J Adv Res Electr Electron Instrum Eng 4:5111–5121. https://doi.org/10.15662/ijareeie.2015.0406029
15. Narang S, Jain K, Saxena M, Arora A (2018) Comparison of face recognition algorithms using Opencv for attendance system. Int J Sci Res Publ 8(2):268–273
16. Kharat D, Kajawe P, Kopnar S, Jinabade N, Wagh J (2019) Face recognition using local binary pattern histogram (LBPH). Int J Innov Res Sci Eng Technol 8(11)
17. Recognition with OpenCV. https://docs.opencv.org/3.4/da/d60/tutorial_face_main.html
18. do Prado KS Face recognition: understanding LBPH algorithm, towards data science. https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b

Recommendation System: Posts Filtering in a Women Empowerment Platform

Devansh Devagya, Harsh Kumar Goel, Harsh Saxena, and Akash Punhani

Abstract The women empowerment platform develops a medium, which would help the women in the society get more job opportunities and an equal stand in the society as compared to men, to get inspired from success of other women and find the right way to work on their weak sections as well in getting the job. The portal also helps to explore the different crime zones in the society by encouraging the women to post about any sort of harm they felt at their workplace or any other place, so that police departments can work on them thus building a safer place for women, where with the pillars of support and safety, women can work with the peace of mind. In this system, the women are allowed to participate in the discussions and help the topic in an open environment. The system also suggests the post the other users based on the content-based recommendation system. The main advantage of using the content-based system is that it does not map the user to user, instead it will map the content to the user. The mapping of the content helps the user to focus on the exactly the post that is required as they may be chance of getting the post liked by the other may not exactly to the other user choice.

Keywords Women empowerment · Recommendation system · Job portal · Crime alert · User-friendly application

1 Introduction

Women empowerment is mandatory for the overall development and growth of the nation. The women empowerment is a very serious issue in India, and they constitute the high proportion of workforce that is low skilled. The low skill in turn means that women will be involved in low-productive and low-paying works. It can be observed that these jobs are highly insecure [1]. It can be observed that from 2005 to 2018, the participation of the women has also decreased by 10% approximately due to socio-economic issues [2]. The various factors like less job, insecurity, and crimes

D. Devagya · H. K. Goel · H. Saxena · A. Punhani (✉)
ABES Engineering College, Ghaziabad, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_6

against are always a point of concern for the society. The various governments have also taken various steps to overcome these issues like "operation durachari" started by Chief Minister of Uttar Pradesh Shri. Yogi Aditya Nath. All these factors have motivated as to design the platform for the women, where they can talk freely, get aware about the various issues going the society, and explore the better opportunities for themselves. The portal will be exclusively for women, and they can speak freely. This will also lead to the special group among themselves that could work on its own by helping each other. The portal will have information related to the various jobs, crimes, frauds, and important information like pandemic-related, and these all information will be posted by the member of this portal. As the members are going to post the information, there will be flood of the posts on the portal. As large volume of post will be there, it will be difficulty for candidate to read all the post; hence, there is the need of specialized recommendation system to capture the needs of the portal. In general cases, the user-to-user mapping is done to perform the collaborative filtering for the recommendation; however, this may lead to some unwanted suggestions that may result wastage of time and diminishes the interest of the users.

The paper has been organized into four sections. In Sect. 1, the details of the portal where this recommendation system is required have been suggested, Sect. 2 discusses the various recommendation systems that are available to use. In Sect. 3, the proposed approach to deal with the recommendation has been stated. In Sect. 4, conclusion has been discussed.

2 Literature Survey

Recommendation filtering plays a crucial role in the latest online services. The basic flow of the recommendation system has been described in Fig. 1. From Fig. 1, it can be observed that initially all the records on the portal are residing on the web server or Internet. A recommender system gets the statistics from the server about the post and reports the same to the user. Based on study [3], the recommendation system can be classified into four categories. The classification has been described in Fig. 2.

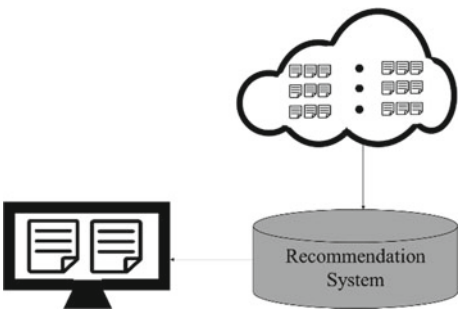


Fig. 1 Describes the component of the recommendation system

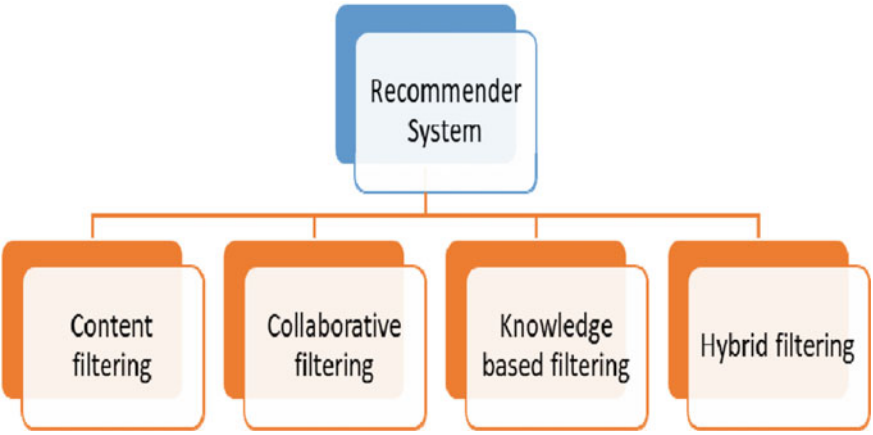


Fig. 2 Describes the various types of recommendation systems [3]

As from Fig. 2, it is clear that there are four different approaches to provide the recommendation.

2.1 Content-Based Model

Content-based model uses the user behavior based on the past activities of the user on the portal or that may be provided by the user in the past. In this technique, the profile contains the details about the likes and dislikes about the specific feature that is recorded. At the same time, the product or item for which the recommendations are to be generated is also profiled based on the feature. While recommending the particular item, the user profile of features is compared with the profile of the features [4, 5].

2.2 Collaborative Filtering (Cf)

In collaborative filtering, instead of creating the items profile individually, the profiles of the users are generated based on the behavior they are performing or selecting the things. The logs of these behaviors are kept and then used to compare the behavior of the users. This mapping of user to user creates the cluster for related categories. This kind of technique is helpful for the recommendations when there is large dataset in consideration. We tend to find similar users and suggest what they want in collaborative filtering. We do not use the item's functionality to recommend it in this form of recommendation system; instead, we group users into clusters of related categories and recommend each user based on the preferences of their cluster [6]. The neural

attention mechanism enables the neighbor model to acquire an adaptive nonlinear weighting function, with the more related users contributing higher weights at the output module. By piling several hops, we later expand the model to a more complex architecture [7]. Sensing and tracking data, such as in mining extraction, environmental sensing over vast fields, or several sensors, financial data, such as financial service organizations that incorporate several financial sources, or electronic transactions and online apps where the emphasis is on customer data have all been subjected to collaborative filtering methods.

2.3 Knowledge-Based Filtering

The knowledge-based filtering works based on knowledge about the underlying problem and identifies the best the solution based upon it. The major advantage of this type of technique is that it helps in exposing the solution that might not be explored due to lack of user profile related to the specific feature [8].

2.4 Hybrid Filtering

It has been observed that the collaborative and content-based filtering have their own constrain, which leads to low performance, and this can be improved if both the techniques are combined together to achieve higher accuracy. This combination can be weighted or any other approach leads to the recommendation system termed as hybrid recommender [8].

Based on the various studies proposed, it can be observed that there is the provision of performing the filtering based on content, collaborative, or knowledge-based or the hybrid of the above. In this article, the posts are tagged using the user choices; however, the study investigated the use of different distance parameter which is not used most oftenly.

3 Proposed Solution

In the proposed approach, the data is collected in the form of post. The posts are having the tags and locations related to the post. The tags are provided by the user that may be appropriate to the content. This information is collected in the database. The data from the database is retrieved, and each of the posts is converted into the vectors X. The various components of the X vector are represented as descried in Fig. 3.

Similarly, the user vector is created based on the requirement filled by her or the post she has liked as described in Fig. 4.

Id	Location	Tag 1	Tag 2	Tag 3	…	…	Tag n
1	Agra	1	0	0	…	…	1
2	Delhi	1	1	1	…	…	1
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
K	Kanpur	1	1	1	…	…	1

Fig. 3 Describes the representation of the post with tags

UId	Location	Tag 1	Tag 2	Tag 3	…	…	Tag n
1	Agra	1	0	0	…	…	1

Fig. 4 Describes the vector of the tags for specific user

In the next step, the cosine similarity is used to compute the idea post based on the vector. The cosine similarity is the kind of the distance metrics that is help full in identifying the similar object. This type of technique is ideally used for computing the similarity in the text [9]. The mathematical representation of the same has been described by Eq. (1) as stated below [10]:

$$\text{Sim}(X, y) = \frac{X \cdot y}{\|X\|\|y\|} \tag{1}$$

Once we got the similarity, the distance is computed using the relation

$$D(X, y) = 1 - \text{Sim}(X, y) \tag{2}$$

This distance is used to identify the K-nearest neighbors. The flowchart representing the same has been described in Fig. 5.

4 Result

In the study, we have compared the two different distance metrics to compare the tags of the different posts. In the proposed approach, the cosine similarity is used to estimate the difference, and similarly, the most popular distance metrics are also

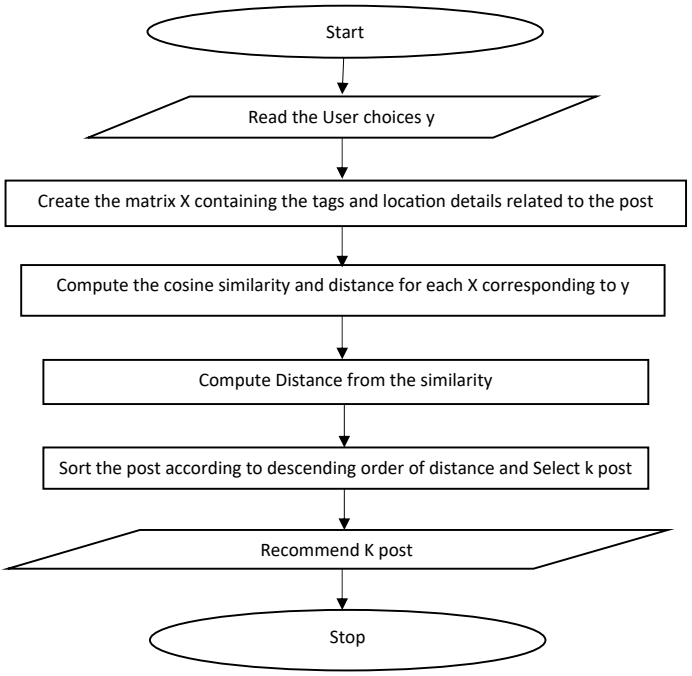


Fig. 5 Describes the flowchart for the proposed recommendation

used for the analysis of the result. In order to analyze the results, three different user tags were generated, which are described in Table 1 along with the best four recommendations based on expert choice.

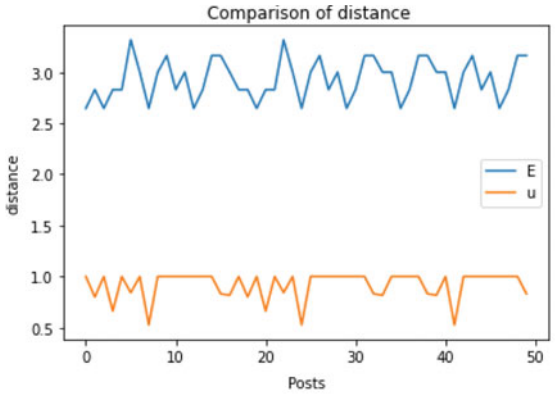
Table 1 Describes the details of the results obtained

User tags	Cosine similarity	Euclidean distance	Expert choice
"Carrier"			
"Work"			
"Carriers"			
"Jobsearch", "jobseekers"	24, 7, 41, 3	0, 46, 41, 35	7, 41, 24, 3
"Workfromhome"			
"Workingathome", "carriers"	33, 39, 16, 32		33,39,16,32
"Jobsearch", "empowerwomen"		2, 19, 21, 4	
"Life", "Healthcare", "COVID-19"	9, 26, 43, 47	36, 47, 30, 9	9,26,43,47

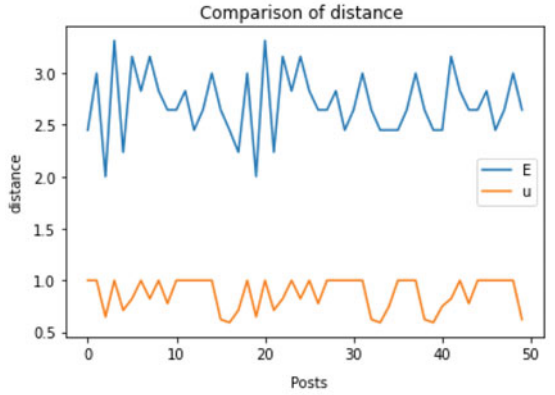
Comparison of the distance evaluation by both the metrics has been represented in Fig. 6a–c that represent the distance computation for the three user tags.

5 Conclusion

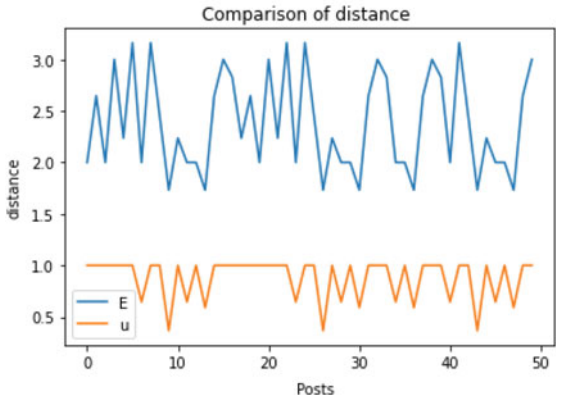
From the results obtained, it can be observed that the recommendation generated using the cosine similarity is better in comparison to the recommendations generated by the Euclidean distances. The results also state that the cosine similarity and Euclidean distance are total difference metrics and have to be selected based on the problem. It can be observed that the K-nearest neighbors estimate is better with cosine similarity. In different studies, the KNN has always being associated with the Euclidean distance. However, as we can observe that here the posts are tagged and for the language processing, it would be better to use the cosine similarity which is not actually the part of KNN. The application of the same has been presented in portals for women empowerment that are necessary for the upliftment of the society, and there is the need for portals. Here, the post of different women will be mapped according to the tags and will help in identifying the better posts for the users.



(a) 'Carrier','Work','Carriers','Jobsearch' 'Jobseekers'



(b)'Workfromhome ','Workingathome','Carriers','Jobsearch','Empowerwomen'



(c) 'Life ','Healthcare','Covid19'

Fig. 6 Describes the comparison of the distance

References

1. Gupta M (2021) Role of NGOs in women empowerment: case studies from Uttarakhand, India. J Enterprising Commun. https://doi.org/10.1108/JEC-04-2020-0066
2. Saraswathy M (2019) Women's day 2019: labour force participation of females is still an Achilles Heel for India. Retrieved from https://www.moneycontrol.com/news/business/economy/womens-day-2019-labour-force-participation-of-females-is-still-an-achilles-heel-for-india-3623061.html
3. Raghuwanshi SK, Pateriya RK (2019) Recommendation Systems: Techniques, challenges, application, and evaluation, pp 151–164. https://doi.org/10.1007/978-981-13-1595-4_12
4. Nikzad-Khasmakhi N, Balafar MA, Reza Feizi-Derakhshi M (2019) The state-of-the-art in expert recommendation systems. Eng Appl Artif Intell 82:126–147. https://doi.org/10.1016/j.engappai.2019.03.020
5. Riyahi M, Sohrabi MK (2020) Providing effective recommendations in discussion groups using a new hybrid recommender system based on implicit ratings and semantic similarity. Electron Commer Res Appl 40:100938. https://doi.org/10.1016/j.elerap.2020.100938
6. Li B, Zhu X, Li R, Zhang C (2015) Rating knowledge sharing in cross-domain collaborative filtering. IEEE Trans Cybern 45(5):1068–1082. https://doi.org/10.1109/TCYB.2014.2343982
7. Ebesu T, Shen B, Fang Y (2018) Collaborative memory network for recommendation systems. In: 41st international ACM SIGIR conference on research and development in information retrieval, SIGIR 2018, pp 515–524. https://doi.org/10.1145/3209978.3209991
8. Laseno FUD, Hendradjaya B (2019) Knowledge-based filtering recommender system to propose design elements of serious game. In: Proceedings of the international conference on electrical engineering and informatics, July 2019, pp 158–163. https://doi.org/10.1109/ICEEI47359.2019.8988797
9. Mohammed SM, Jacksi K, Zeebaree SRM (2021) A state-of-the-art survey on semantic similarity for document clustering using GloVe and density-based algorithms. Indonesian J Electr Eng Comput Sci 22(1):552. https://doi.org/10.11591/ijeecs.v22.i1.pp552-562
10. Muflikhah L, Baharudin B (2009) Document clustering using concept space and cosine similarity measurement. In: ICCTD 2009—2009 international conference on computer technology and development, vol 1, pp 58–62. https://doi.org/10.1109/ICCTD.2009.206

Evoting Using Blockchain Technology

Amrita Jyoti, Rashmi Mishra, Rupa Rani, and Ravi Kalra

Abstract Voting is the fundamental right of each and every citizen in a democracy. Perhaps, it is the very basis of a democracy. Democracy means a government for the people, of the people, and by the people. This means many rights such as freedom of speech, right to information, to name a few. But the right that defines a democracy is voting. It means that those who are willing to form a government and serve the people stand as candidates, and the people are allowed to choose, who they wish to be served from (i.e., the government). Looking at how imperative voting is, it is also essential that all citizens are able to cast their votes during the election. Now here comes the problem, especially in countries like India. Some people are not available in the city at the time of voting, they may be traveling, etc.; for the others, the voting center might be very far away from their houses. And in India, with the vast multitude of population, a large fraction of people is not even able to receive their voter ID cards (Barański et al. in Appl Sci 10:7606, 2020 [1]). Also, some people may be handicapped and hence are not able to vote. There is a one clear solution to all this, "Evoting," a concept that will be elaborated in this paper, stating how it is done, what technology it uses, and its advantages over the currently used voting methods.

Keywords Voting · Blockchain · Evoting · Ballots · Electronic voting machine · Public distributed ledger · Mining · Miners · Transaction · Immutable · Proof of work · Cryptocurrency · Encryption · Democracy · Election · Ganache · Truffle · JavaScript · MetaMask · NPM · Solidity

A. Jyoti (✉)
Ajay Kumar Garg Engineering College, Ghaziabad, U.P, India

R. Mishra · R. Rani · R. Kalra
Krishna Engineering College, Mohan Nagar, Ghaziabad, U.P, India
e-mail: rupa.rani@lloydcollege.in

R. Kalra
e-mail: ravi.kalra@lloydcollege.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_7

1 Introduction

The problems stated in the abstract are prevalent all over the world, with the system that is used for voting. Currently, voting is carried out in 2 ways:

1.1 Paper Ballots

Here, the voter needs to select/mark the party he/she wants to vote for on a piece of paper and then put the paper in a box. This results in lots of paper usage, thereby making this method very less eco-friendly due to the carbon footprint it has. Also, since the votes are cast on paper, the process of counting takes place manually [2]. This has some problems associated with it.

Manual Errors: The process of counting such a large number of votes is tedious, and the people can get lethargic while counting the votes which can result in the officials' losing track of the counting. So, the counting loses its reliability and credibility, which is so very important in voting.

Cost Of Voting: Manual counting means the counting will be done by humans. And this human labor will be needed to be paid some wage. This may sound a little easy in countries like India and China (where the cost of labor is very less), but in countries like the USA the cost of labor is extremely high; this does not seem feasible. With the labor comes a cost, and this cost becomes a disadvantage in ballot-based Evoting.

1.2 Less Eco-Friendly

Also, since this method of voting uses paper, it has lots of carbon footprint associated with it, directly proportional to the paper produced in making the ballots, and then resource consumption in printing the ballot on the paper.

1.3 Time Taken to Announce Results

Now, since the counting is manual (which it is in most cases), the time taken to announce the results is also very high. And, even if we consider the time taken to count the votes using machines, it would still be considerably high.

2 EVM Machines

EVM is an initialism for electronic voting machine. In this case, the voter needs to go to the voting center and press the button of the party of the voter's interest, and the vote is cast electronically [3]. In this case also, the following problems arise:

2.1 VVPAT

VVPAT is an initialism for Voter-Verified Paper Audit Trail. This is a system in which the cast vote of the voter is printed on a piece of paper, which is displayed through a transparent screen to the voter. This allows the voter to verify whether his/her vote has been correctly recorded. Now, this VVPAT system (which is mostly employed with the EVM, except for some rare cases) brings with it certain issues. First, it brings back the carbon footprint that was also an associated problem with paper ballot voting system. Second, it necessitates the integration of a printing device with the display system of the EVM, thereby increasing the consumption of electricity and other resources [4].

2.2 Increased Cost

The EVMs are operated electronically and also require other devices such as printers integrated with them. This leads to costs such as electricity, printer inks, paper procurement for the VVPAT paper slips, and also a very big cost which is called the maintenance cost of the machines. Also, keeping the machines secure is a cost (the storage cost).

2.3 Danger of Tampering

In the recent elections, there have been claims by some political parties that the EVMs can be tampered and hacked. If the EVM is hacked by any means, there would be no credibility and reliability of the voting conducted; hence, the essence of a democracy is lost.

2.4 *Time for Result*

One more problem with EVM voting is the tedious process of vote counting. Though electronic, this method also takes some time in vote counting as all the EVMs from the different locations would be collected at one central place, and then the results of the different EVMs would be added up. Now, one major issue arises at this point. If, at the time of vote counting, there is a dispute, then the EVM total would have to be matched with the VVPAT paper total, and for this, the VVPAT paper ballots would have to be counted, thereby resulting in the EVM voting degrading to the cumbersome paper ballot voting task.

3 What is Blockchain

A blockchain is a collection of blocks, each of which stores some information. A block in a blockchain takes its reference from the previous block in the chain to verify and perform changes to the data in the previous block that are reflected in the current block. Each such change in a blockchain, wherein a new block is added to the chain, is called a transaction. Now, this chain of blocks is distributed to everyone in the world and hence is known as a public distributed ledger (Fig. 1).

Now, there are some agents who add new blocks to a blockchain, called "miners." These miners have to solve a complex mathematical problem to add a block to the chain [6]. This is called "proof of work." After the miner has solved the problem, the block gets added to the chain (this is called "mining"), and the miner gets a reward

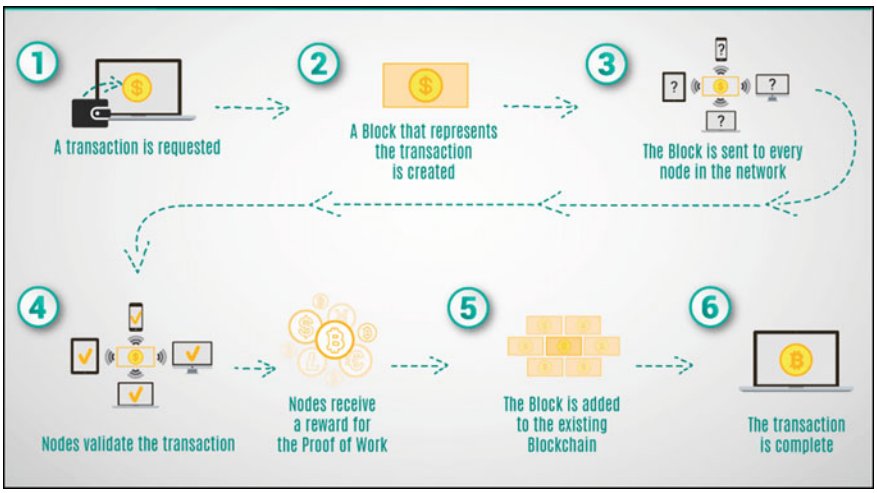


Fig. 1 How blockchain works [5]

in the form of cryptocurrency. All cryptocurrencies (e.g., Bitcoin, Ethereum, etc.) work on the concept of blockchain.

4 How is Blockchain Used in Evoting

Let us see how the above-discussed blockchain technology is used in Evoting. This concept is implemented in a different type of web called the distributed web (d-web). The difference between normal web and d-web is that the web is centralized in nature, whereas the d-web is distributed in nature. In the case of the web (which is currently being used by everyone), we follow the client–server paradigm. In this case, the user is the client and the entity where the actual data stored is called the server. When the client requires any information, a request is sent to an entity in the network that contains that information (i.e., server). The information may be anything such as a web page, a pdf, a document, and an image. Now, if the server has the information, it sends that information to the client [7]. But in this case, there is always the risk of central failure, as all the information resides in the server, and failure of the server would mean failure of the entire system. Enter the d-web. In d-web, there is no central entity governing everything. But anything and everything a user may require is distributed among all the nodes in the network. This acts more like a peer-to-peer network. So now, if a node requires some information, then its request can be served by all the other nodes in the network. These other nodes have either full copies of the document or parts of it (in this case, the client's request can be fulfilled by multiple nodes having the copies). This enables the big, combined computing power of the different nodes to be utilized in the most optimum way [8]. It also eliminates the need of depending on a central authority for running the show. Also, we will not have to give ownership of any document to any entity (as in case of a cloud). And the publication process becomes much easier. There is no need to own a server in the d-web. Taking all the above facts into consideration, the process of Evoting would be carried out in this distributed web. The voter may access the online polling booth in two ways, either through an application that accesses the d-web (i.e., a d-app or a distributed application) or through their web browsers using an interface to access the d-web (e.g., MetaMask, etc.). Now, at the side of the organizer, there is the task of maintaining the database of the eligible voters, the creation of an election, and the maintenance of the vote blockchain (a blockchain that is a record of the votes cast).

Before the election, the election organizer will create an election. On the election day, the voter will go to the voting app (in case of d-app usage) or would log in to the voting portal (in case of web browser usage). For verifying the identity, the voter would need to enter credentials such as mobile number and Aadhaar number. Then, an one-time password (OTP) would be sent to the voter's mobile. Upon validating the OTP, the voter will be directed to the Voter Registration Page. Here, upon requesting for registration, the voter credentials would be sent to the organizer, who will verify whether the voter is eligible to vote from the voters' database. If the voter is found to be eligible, the voter's details are added to voters' blockchain (a blockchain storing the

registered voters' details and also the votes cast by them) and a smart contract will be issued to the voter in the form of an electronic ballot. The voter would then cast his/her vote on this electronic ballot and submit the vote. On pressing the submit button, another OTP would be sent to the voter's mobile number to confirm submission of the vote. The voter will enter this OTP and log out of the voting page. This vote would also be added to the voters' blockchain. Since the voters' blockchain is a public distributed ledger, the voter can audit the votes any time, while also maintaining the privacy of other voters. Hence, fair results are ensured. Also, the results can be immediately declared after the voting with no manual counting process required, thereby completely eliminating the chances of errors due to counting. This process has no carbon footprint as no paper ballots are used in the voting process. Also, there is no cost of running the heavy machinery and no consumption of expensive resources and the costs associated with them. And most importantly, this method eliminates the need of physically going to a polling booth for voting. All that is needed is a stable Internet connection, and the job is done, sitting at home, hassle-free [9].

5 Phases of Evoting Using Blockchain (Explained with the Help of Flowcharts)

The entire process of Evoting using the blockchain technology can be divided into the following two phases.

5.1 Registration Phase

In this phase, the following events occur: The election organizer will create an election on the election app. The voter will download an Evoting app or would use the voting portal over the distributed web through web browser. He/she will log in using credentials like mobile number and Aadhaar card number on the app for registration. An one-time password (OTP) would be received on the voter's mobile number. Upon entering the correct OTP, the voter will be directed to the registration page. The voter would request for registration. Now, the election authority would validate whether the voter exists in the voters' database. If yes, the voter's details would be added to the voters' blockchain and a smart contract will be issued to the voter in the form of a ballot, on which the voter can cast his/her vote on the election day, using their app (see Fig. 2).

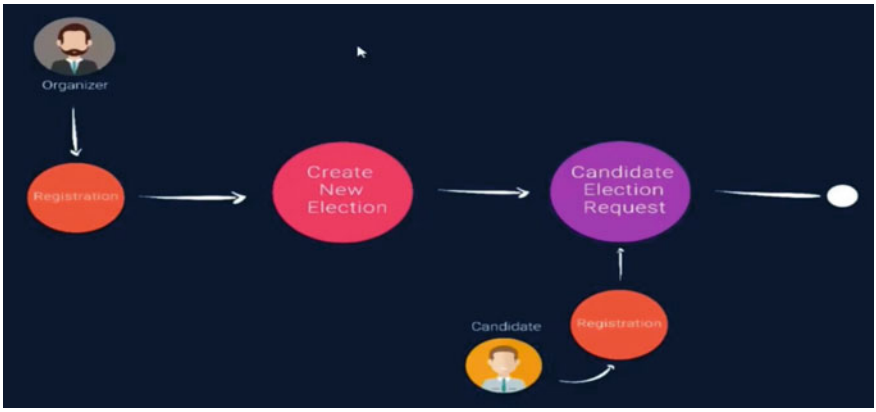


Fig. 2 Registration phase in evoting [10]

6 Voting Phase

First, this phase takes place on the day of the election. Secondly, the voter opens his/her app and requests to vote. Then, the election authority checks whether the user is valid and if this is their first vote. If the above details are verified, the voter is allowed to vote. The voter then casts his/her vote. Then, an one-time password (OTP) is sent to the voter's phone for confirming the vote submission. On successfully entering the OTP, the vote gets added to the voters' blockchain (see Fig. 3).

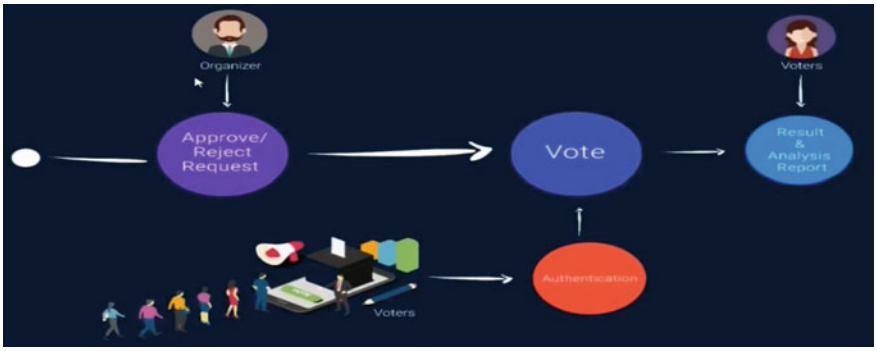


Fig. 3 Voting phase in evoting [10]

7 Tools and Technologies Used

This system of Evoting can be implemented with the following tools and technologies (as explained in this section in the form of a classification of hardware and software requirements):

7.1 Hardware Requirements

For implementation of the proposed voting method, the organizer of the election would need a personal computer (Windows 7 onward). It needs to be a 64-bit machine. As far as the voter side is concerned, the voter can vote in two ways. Using a personal computer (laptop, desktop, etc., Windows 7 onward for Microsoft users and all versions of operating systems for Mac users; the machine needs to be 64 bits) either in case of voting using the web portal [11] or in case the voter wishes to use the d-app, he/she may use a smartphone (Android or iOS both would work).

7.2 Software Requirements

This section only concerns with the software requirements of the election organizers. Since this voting method would be implemented on the distributed web, the organizer would require some advanced tools and technologies to create and host the elections, and accept votes in the blockchain, keeping them secure and maintaining the reliability and credibility of the elections.

The following are the tools required to conduct election using the blockchain technology.

NPM: NPM is an initialism for Node Package Manager. NPM can be used in two ways, as an online platform and as a command line tool. As an online platform, NPM allows one to publish and share tools made using the JavaScript language [12]. These tools can be used in browsers (i.e., in front end), servers (i.e., at the back end), or even the command line. As a command line tool, NPM enables users to interact with the aforementioned online platforms. So, one can do the following things with NPM as a command line tool:

It enables one to **install and uninstall packages**. A package is a tool someone created and uploaded to the NPM online platform. One can think of a package as a building block of an application or an even bigger package.

NPM can also be used for **version management**. Every package has a version. As the package grows and improves, the version of the package also changes. With NPM, one can easily keep the packages in one's project up to date. It even lets one switch to another package version, whenever one wants.

It is used for **dependency management**. Many packages that people create are built on top of other packages. Dependencies are packages on top of which other packages are built. Instead of the developer searching for each and every dependency and installing them one by one, he/she can use a single command to install a package with all its dependencies. This means that NPM does all the hard work for the developer.

Truffle framework: Truffle framework is a very powerful tool that is used to work with Ethereum smart contracts. It is used for deploying, compilation, and linking of smart contracts, manages networks and packages, and also, provides a testing platform for automated contracts, etc.

Ganache: Ganache was previously known as testrpc. It can be used in two ways, as a command line and as a user interface. It uses a virtual blockchain, and then, it establishes ten standard Ethereum addresses with all the blocks in the chain along with private keys. Ganache then preloads them with hundred ether each (theses are simulated ether). There is no mining required in Ganache. Instead, the task is made much easier with Ganache automatically confirming every transaction. It is convenient to use with a vast variety of operating systems like Windows, Linux, and Mac.

MetaMask: MetaMask is an open-source tool. It is extremely user-friendly tool. It has a graphical user interface (GUI) and performs all transactions in the Ethereum cryptocurrency. The common web browsers (like Google Chrome, Mozilla Firefox, and the like) are not capable of running or one can say, accessing the distributed web (d-web) on their own.

They need a support for that, and MetaMask is that support. MetaMask allows Ethereal DApps to run on one's normal system browser without the user node having the need to become a complete Ethereum node. One can say that MetaMask is a bridge between the normal system browser and the distributed web.

Solidity: Solidity is a language that uses JavaScript style syntax for contracts. It is used to generate EVM (here, EVM stands for the Ethereum Voting Machine) machine level code and then further convert it into simple instructions. It works with four value types. These are address, integer, string, and Boolean. Solidity uses the same operators as those used in JavaScript. It is a high-level language.

8 Advantages of Evoting Using Blockchain Over Other Systems

The advantages of Evoting using blockchain are:

8.1 Auditability

The voter can audit all the ballots in the chain, while maintaining the other voters' privacy. This becomes possible because the voters' blockchain is a public distributed ledger, granting access to all the voters in the voters' list.

8.2 Transparency

Maximum people, in the current voting system, think that their vote is not considered due to some fraudulent practices that may be taking place at the time of vote count. But this problem is tackled properly by the method proposed in this paper. The voters' blockchain is transparent, and the voter comes to know whether or not his/her vote is counted. All this becomes possible due to the distributed nature of the blockchain.

8.3 Immutability

Now, in everything that involves the web, there is always a risk of changing the data contained in the data repository. Switch to blockchain. The blockchain is immutable.

8.4 No Risk of Hackin

In the proposed system, wherein everything is shared in public, there is always a risk of hacking and altering of the data. The voters' blockchain cannot be altered by hackers owing to the encryption applied to every block in the chain. Also, if at all the data of one block is changed, the hacker will need to change all the previous blocks in the chain.

8.5 Secure and Hassle-Free

The main problem with offline voting systems was that the voter had to physically go to the polling booth for casting his/her vote. But now, one can sit in the comforts of one's home or anywhere in the world and vote for one's desired candidate. Also, no security concerns owing to the blockchain technology. Just a stable Internet connection leads to an extremely secure and hassle-free method of voting.

8.6 No Cost of Resources

There is no need for the organizer to procure expensive machines or paper and pay for the heavy amount of electricity and other resources consumed for voting. All the organizer needs are a voting d-app, a setup to host the voters' database and the voters' blockchain, and an OTP generation equipment. And all the voter needs are a mobile app and an Internet connection to vote.

8.7 No Need to Go to Polling Booths

The most concerning issue with offline voting systems was that the voter had to physically go to the polling booth for casting his/her vote. But now, one can sit in the comforts of one's home or anywhere in the world and vote for one's desired candidate. The voter can vote from any part of the world.

8.8 Reliability

Since every vote is stored in the voters' blockchain, every time a voter casts his/her vote, the d-app checks the voters' blockchain for a vote cast by the voter. If no vote from the particular voter is found, it means that the voter is voting for the first time. Only then is the voter allowed to vote. Otherwise, he/she is rejected. Hence, every voter can vote only once.

8.9 No Fraudulent Practices

Since the blocks in the blockchain are encrypted, hence no one can tamper them. Also, the blocks are distributed in public. So, there cannot be any manipulation in the counting of the votes. In short, there is no fraud in voting.

8.10 Immediate Result Declaration

The voting results are announced immediately after voting is completed. There is no requirement of manual counting (as in the case of paper ballot voting) or even machine counting (as in the case of EVMs). So, it is very time saving.

8.11 Environment-Friendly

In this system of voting, no paper is used (which are used in both, paper ballot voting and EVMs). Also, there is no electricity consumption, which is very high in the case of voting using EVMs. Therefore, this method is more environment-friendly, as carbon footprint is negligible. Since there is no need to go physically to the polling booth, there is no use of vehicles, no traffic jams at the polling centers, and hence no fuel consumption (something that contributed largely to environmental pollution).

8.12 No Labor Cost for Voting

Since no manual counting of votes takes place, and also, no vote totaling is done (as in the case of EVMs where the votes of the individual machines are totaled), there is no dispute in the vote count (as happens sometimes with EVMs), and there is no need of any human labor to count the votes [13].

9 Conclusion

Hence, we have a hassle-free system of voting, more secure than anything the world could ever imagine. Extremely, user-friendly voting environment, no problems of going to any polling booths physically, more transparent than ever, no frauds, voting from the comforts of your home, or even while traveling to some other city or abroad, what else does a voter want. There cannot be an easier way for a voter to exercise his/her rights. This practically feasible voting system comes as a boon to all handicapped people who face extreme difficulties in voting and prefer not to vote. Also, in this time of the COVID-19 pandemic, where it is very unsafe to step out of one's house, this system of electronic voting comes to a voter's rescue. Apart from this, Evoting also provides immediate election results. This paper provides a deep insight to the basic concepts required to understand Evoting using the blockchain technology, and also, how it can actually be carried out on a large scale [14]. Apart from a government election, it can also be used to carry out elections in offices, colleges, and any places requiring choosing representatives through an election.

References

1. Barański S, Szymański J, Sobecki A, Gil D, Mora H (2020) Practical I-voting on stellar blockchain. Appl Sci 10:7606. https://doi.org/10.3390/app10217606

2. Lalam N, Nithinn MS, Jebakumar R (2020). BEVS—blockchain based E-voting system. Int J Adv Sci Technol 29(9s):6241–6249. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/20208
3. Wang B, Sun J, He Y, Pang D, Lu N (2018) Large-scale election based on blockchain. Proc Comput Sci 129:234–237
4. Yli-Huumo J, Ko D, Choi S, Park S, Smolander K (2016) Where is current research on blockchain technology?—a systematic review. PLOS ONE 11(10)
5. https://datascience.foundation/img/pdf_images/fig_2_blockchain_architecture_diagram.png
6. Barnes A, Brake C, Perry T (2018) Digital voting with the use of blockchain technology. Available https://www.economist.com/sites/default/files/plymouth.pdf [Nov 20, 2018]
7. Hjálmarsson FÞ, Hreiðarsson GK, Hamdaqa M, Hjálmtýsson G (2018) Blockchain-based E-voting system In: 2018 IEEE 11th international conference on cloud computing (CLOUD), San Francisco, CA, pp 983–986
8. Kshetri N, Voas J (2018);Blockchain enabled E- voting. www.computer.org/software
9. Çabuk UC, Adıgüzel E, Karaarslan E (2018); A survey on feasibility and suitability of blockchain techniques for the E-voting systems. Int J Adv Res Comput Commun Eng
10. https://youtu.be/WLwfCgLHK5Y
11. Cruz JP (a), Kaji Y(b) (2017) E-voting system based on the bitcoin protocol and blind signatures. IPSJ Trans Mathem Model Appl 10(1):14–22
12. Madise Ü, Martens T (2006) E-voting in Estonia 2005. In: The first practice of countrywide binding Internet voting in the world. Electronic Voting, vol 86
13. Green T, Sarrasin O, Baur R et al (2016) From stigmatized immigrants to radical right voting: a multilevel study on the role of threat and contact. Polit Psychol 37(4):1–22
14. Burton C, Culnane C, Schneider S (2016) vVote: verifiable electronic voting in practice. IEEE Secur Priv 14(4):64–73

Conversion of Sign Language into Devanagari Text Using CNN

Vimal Gaur, Rinky Dwivedi, Pankhuri, and Shivam Dhar

Abstract Sign language is an increasingly valuable interactive medium for both deaf and hearing disabled individuals. These languages are considered as a group of predefined languages that use the visual–manual mode to communicate knowledge. These languages are considered to be unitary for deaf people. This paper proposes a useful hand motion picture recognition design method to provide a simple way to interact with the visually impaired and the average person. This is accomplished by considering Devanagari Sign Language finger-spelling comprehension in real time. In this paper, authors have created dataset of Devanagari Sign Language by capturing the gestures through the Web camera. From these captured images, we found the region of interest (ROI). Convolutional neural network has been used for classification as it eases the process of feature creation. For training this network, a hand signal has been used as input and the resultant output produces character when a suitable match is found with an accuracy of 98% and testing accuracy of 99.59%.

Keywords Deep learning · Convolutional neural network (CNN) · Devanagari sign language · Region of interest (ROI)

1 Introduction

India is the second-most populous nation in the world with over a population of billion people. More than a million people are now suffering from hearing problems and vision loss. According to a survey, it has been calculated that one percent of people are deaf and twelve percent have hearing problems. With the introduction of artificially intelligent computers combined with the availability of big data

V. Gaur (✉)
Maharaja Surajmal Institute of Technology, JanakPuri, Delhi 110058, India
e-mail: vimalgaur@msit.in

R. Dwivedi · Pankhuri · S. Dhar
Maharaja Surajmal Institute of Technology, C-4 Janakpuri, New Delhi 110058, India
e-mail: rinkydwivedi@msit.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_8

and vast computing tools, there has been a massive rise in healthcare, robots, automated self-driving cars, and human–computer interface (HCI) [1]. HCI is used in augmented reality systems, face recognition systems, and hand gesture recognition systems. This paper aims to identify different alphabets of the Devanagari family which is surely a part of the human–computer interface. The main aim is to deal with the difficulties faced by deaf and dumb people in communicating with people in their day-to-day life. For effective communication, some signs are required. In this paper, these signs have been gathered together to construct a sign language that differs from one region to another [2, 3]. Nearly, 466 million individuals around the globe experience the problem of deafness (upwards of 5% of the global population). Of those 34 million include children, as per the World Health Organization (WHO) [4]. These estimates are expected to rise by more than 500 million by the year 2050 [9]. Moreover, in several cases, severe deafness disease is a global clumped in lower and middle-income nations. There has been a lot of work in this area in the past using sensors (like a glove sensor) and other image processing techniques (like edge detection, Hough transform,), but these have not yielded satisfactory results. However, with new deep learning techniques such as CNN, productivity in this area has skyrocketed, and opening up many new possibilities in the future [5]. People with hearing disabilities remain back at Internet conversations, workplaces, information sessions, and classes. They mostly use text messages to communicate with others, which is not the ideal tool. Only with the growing acceptance of telemedicine services, deaf individuals can interact spontaneously via their healthcare systems, friends and coworkers, irrespective of how much the other person is familiar with the sign language. Hand gesture recognition systems can be generally grouped into two categories: the first one is the hardware-based system and the second one is the technical vision system. In hardware-based systems, individual needs certain instruments to extract the characteristics that describe the signal of the hand gesture. The cyber glove is an extraction device that is used for extracting the characteristics like orientation, gestures, and color of the hands. It is commonly seen in the recognition of sign language. Vision-based devices use optical imaging systems that extract the lines and interpret the signal. The model demonstrated in this paper is a vision-based solution where no special device has been worn by the consumer. This paper has been structured as: Sect. 2 presents literature review. Problem statement has been presented in Sect. 3. Based upon this problem statement, objectives of research have been formulated and described in Sect. 4. Section 5 presents methodology. Section 6 presents results and discussions, and Sect. 7 concludes and presents future scope.

2 Literature Review

As per the World Health Organization (WHO), the estimates of having to hear people have finally hit 400 million. It is due to this fact that current findings have been intensified to make it possible for people with disabilities to interact. In recent years, several researches have been carried out. Few of them have been listed as below:

In [1], the authors have created smartphone SLR software for the identification of Indian Sign Language (ISL) gestures. In this research work, they have used the selfie mode continuous sign language capture method and CNN model for image recognition. In [2], the researchers have proposed a sign language recognition system that converts sign language in real time, thus allowing people who are not familiar with sign language to interact effectively with having to hear people. American Sign Language has been taken as input, and convolutional neural network has been trained on the dataset by Massey University, Institute of Information and Mathematical Sciences (2011). After the network has been trained, the network model and network weights have been stored for the real-time recognition of sign languages. Also, they have used certain frames for the determination of skin and for the hand gesture determination they used a convex hull algorithm. In [3], the researchers have focused on creating vision-based software that includes the text-based interception of hand gestures, thereby promoting contact between signatory and non-signatory. The dataset used by them is American Sign Language (ASL). The proposed model accepts the video frames as input and removes spatial and temporal information from them. In this research, they have used CNN to identify the spatial features and RNN to train on the temporal features. In [4], the authors have used the strategies for segmentation of images and identification of features. They have employed FAST and SURF algorithms to establish a relationship between image segmentation and object detection. The system created by them goes through multiple phases, such as capturing the data using the KINECT sensor, performing the segmentation of images, feature detection and extraction of ROI. The K-nearest neighbors (KNN) algorithm has been used to distinguish images. Text-to-Speech (TTS) conversion has been performed for the audio output. In [5], an Indian Sign Language translator has been created using a convolutional neural network algorithm by the researchers to identify the 26 letters of the Indian Sign Language into their corresponding alphabetic characters by taking a real-time representation of the gesture and translating that to its identical. In the implementation, a database has been generated using different image preprocessing methods to make the database available for the extraction of a feature. After the features were extracted, the images were sent to the CNN model using the Python program. In [6], a sign language finger-spelling alphabet detection system has been proposed using different image processing methods, supervised machine learning, and deep learning. The histogram of oriented gradients (HOG) and local binary pattern (LBP) characteristics within each symbol has been derived from the input image, and afterward, the multiclass support vector machines (SVMs) were used to train such collected features, and the end-to-end convolutional neural network (CNN) architecture was employed in the training set for the comparison. A Bangla Sign Language to text translator device has been made [7] by the researchers through customized region of interest (ROI) segmentation and convolutional neural network (CNN). They used customized ROI segmentation to allow the system person to change the pre-loaded boundary box on the display screen to the deaf person's hand area and thus only the area inside the video frame would be forwarded to the CNN prediction model. The model has also been incorporated with the ARM cortexA53 embedded Raspberry Pi, which provided durability and portability to the

system. In [8], the authors of the paper have designed an American Sign Language translator app Using OpenCV. They have used a ResNet-34 CNN Classifier for the classification of American Sign Language (ASL) hand gestures. "Firstly this application extracts the signs from the input video, this video then converted into frames, later ResNet-34 CNN classifier has been used for classifying frames and finally the text corresponding to the sign is displayed as an output on the screen". In [9], the researchers carried out a consistent study of the numerous methods used to translate sign language to text/speech form. Later, they have employed the best possible method to create an android app which can convert the ASL signs to text or speech in real-time.

After doing a thorough analysis on these studies, it has been analyzed that there is a requirement of real-time software for detecting hand gestures so that they can be converted into Devanagari script. Table 1 shows the methodologies that have been adopted by different researchers.

Table 1 Methodologies by different researchers

S. no	Author	Dataset	Accuracy (%)	Algorithm
1	Rao et al.	Real-time images	92.88	CNN
2	Taskiran et al.	Massey University, Institute of information and mathematical sciences	98.05	CNN
3	Bantupalli et al.	American sign language dataset	Created a vision-based application for sign language to text conversion	CNN(spatial features) RNN (temporal features)
4	Andewale et al.	Images are collected using Kinect sensor	78	Unsupervised machine learning algorithms
5	Intwala et al.	Real-time images	87.69	CNN
6	Kania et al.	Exeter University Dataset and Massey University web pages	93.3	Snapshot learning model
7	Khan et al.	Custom image dataset	94	Implemented in Raspberry Pi
8	Kurhekar et al.	Real-time images	78.5	A ResNet-34 CNN classifier
9	Tiku et al.	Real-time images	98.82	SVM

3 Problem Statement

In this paper, a real-time software has been developed for detecting hand gestures and converting them into a Devanagari script, using a deep learning approach (CNN). The results are required to be compiled using 25 alphabets of the Devanagari script.

4 Objectives of Research

- To generate a dataset that is skin tone independent using a webcam.
- To apply appropriate image processing and preprocessing techniques for achieving better accuracy and to obtain region of interest (ROI).
- To design the model and building CNN to train raw images and achieve optimal precision.

5 Methodology

5.1 Dataset Generation

Due to non-availability of datasets on Devanagari script, we instigated to create the datasets with 200 images from different angles. For the creation of the dataset, we used the webcam of personal computer. The webcam used for data gathering purposes is of 2.1 Megapixel and relative illumination is greater than 40%.

Steps to create dataset: OpenCV module has been used to produce dataset. Nearly 600 images of each alphabet of Devanagari script have been captured for training purposes and 200 images for testing purposes. When the frame of the camera opens, it displays everything in RGB values, but a specific region of interest (ROI) has been created inside that frame only, in the form of a small rectangle, which has videos displayed in the form of adaptive Gaussian filter. The small rectangle in our frame converts RGB into grayscale and finally applies an adaptive Gaussian filter in real-time [10, 11].

5.2 Gesture Classification

We have used one algorithm for predicting the final symbol displayed by the user.

- Applying an adaptive Gaussian filter and threshold to the frame, take using OpenCV, to extract features and then result in the processed image.
- Image after processing is passed through CNN model for prediction, and if a letter is recognized by the system, it is printed.

- In case more than one letter shows a similar result, we tried to increment the number of distinct positions

A. Layer 1: CNN model:

- First Convolutional Layer: The input picture is first passed through the convolutional layer. This layer uses 32 filter weights.
- First Pooling Layer: The pictures are down sampled using max pooling. The resultant output is a 2 * 2 array.
- Second convolutional layer: This array is then processed using 32 filter weights. The output of this layer is reshaped to an array of 30*30*32 = 28,800.
- Second Pooling Layer: This processed image is then down sampled again and reduced into images with less resolution.
- First Densely Connected Layer: These less-resolution images along with an input array of 28,800 are send to first densely connected layer having 128 neurons. The output of this will be used in the second densely connected layer. A dropout layer with a value of 0.5 has been added to avoid the overfitting of the model.

B. Layer 2: Activation Function:

- Rectifies layer unit has been used in every layer (convolutional as well as fully connected neurons). For each input pixel, max(x, 0) is calculated by ReLu. This adds non-linearity to the formula and trains the model on more complicated features. Due to this, vanishing gradient problems is also eradicated and the training process also fastens.

Layer 3: Pooling Layer: Max pooling layer has been applied to input image with a pool size of (2, 2) with ReLu activation function. This lessens the number of parameters and ultimately leading to reducing computation cost and overfitting of the model.

Layer 4: Dropout Layer: The problem of overfitting has been resolved by tuning the weights of the network. So, in dropout layer a random set of activations sets these weights to zero. Thereby network helps in providing accurate classification even if some activations are dropped out.

Layer 5: Optimizer: For updating the model in accordance with the output of loss functions, an Adam optimizer has been used. It adds benefits to the system by combining extensions of two stochastic gradient descent algorithms, namely adaptive algorithm (ADA GRAD) and root mean square propagation (RMSProp).

5.3 Training and Testing

Initially, the input images in the colored form have been converted into grayscale images and then passed under an adaptive Gaussian filter for removing unwanted noise. The complete dataset is divided into 70% training and 30% testing. The input images after preprocessing have been fed to the model for testing and training. The prediction layer estimates how likely the image will fall in which category. So the output is normalized between 0 and 1 such that the sum of values in every class sums to one. We have achieved this using the softmax function. The result of the prediction layer varies as compared with the actual result. That is why the model has been trained under labeled data. Classification uses cross-entropy as performance measurement, which is a continuous function with assigns positive values to those which are not the same as the labeled value and zero value to those which are the same as the label. We adjusted the weights of neural networks such that the cross-entropy is minimized to zero. An in-built function to calculate cross-entropy is already present in TensorFlow. After identifying the cross-entropy function, we have optimized it with Adam optimizer.

6 Results and Discussions

While training the dataset without augmentation, the CNN model achieved high training accuracy up to 98% and 99.59% testing accuracy, demonstrating real-time environment results. As discussed, region of interest has been demonstrated in Fig. 1.

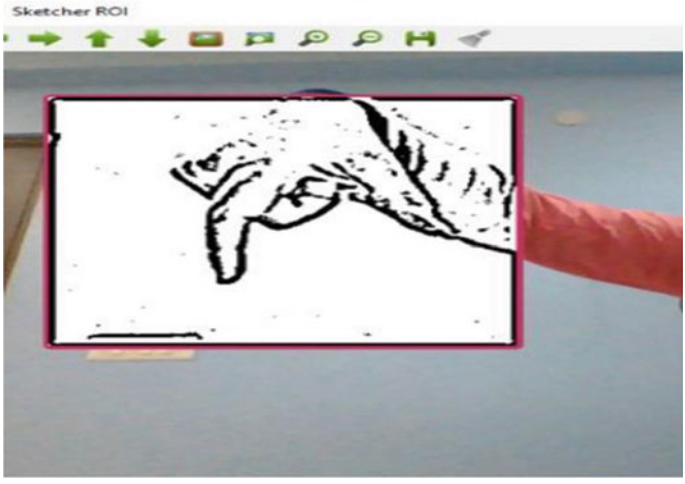


Fig. 1 Demonstration of region of interest (ROI)

Figures 2 and 3 shows the snippets of the model. As shown, accuracy varies from 95.13 to 99.59% when the number of epochs increase from 1 to 5.

The difference in accuracy is because hand gestures in real-time are not in the same position. Also, the pixel position and shapes of the images are different. Additionally, the background is scrambled. The comparison of accuracy of the CNN model with the current state-of-the-art have been shown in Table 2.

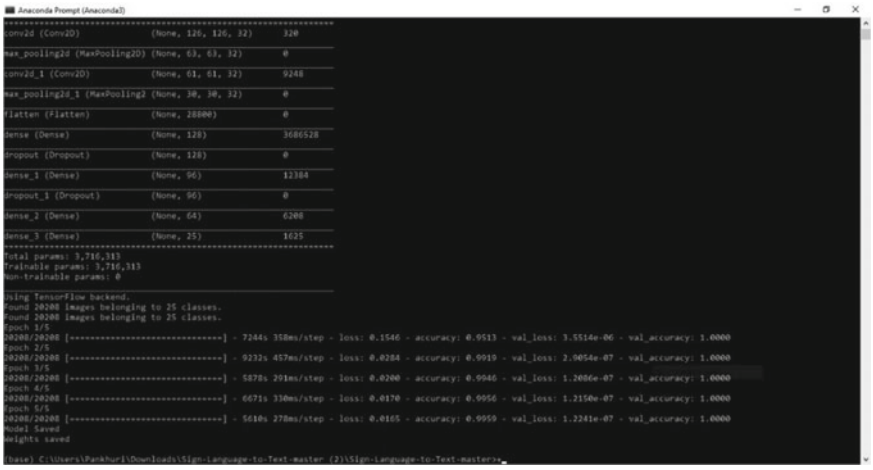


Fig. 2 Snippet 1 of the model

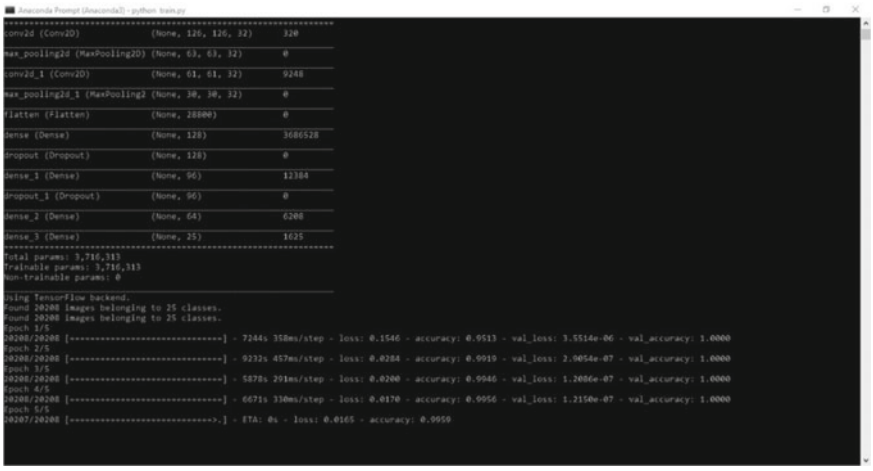


Fig. 3 Snippet 2 of the model

Table 2 The accuracy comparison of the CNN model with the current state-of-the-art methods

Method	Accuracy (%)
CNN with Gaussian filter	99.59
CNN [1]	92.88
CNN [2]	98.05

7 Conclusion and Future Scope

The study shows that Devanagari script used in India, Nepal, and Tibet includes 25 different letters and is written from left to right. Hence, 600 different images were collected as training set and 200 for testing test. From this collection of images, a database of twenty-five different signs were used. These set of images are subjected to batch segmentation, detection of each alphabet is done, and region of interest is extracted from specific bounding box. The combination of adaptive Gaussian filter and CNN showed that the gesture classification could determine the best matched feature. This is achieved by comparing it with existing database. The accuracy percentage while training the model comes out to be 98%, and testing accuracy has been calculated as 99.59%. We have achieved higher accuracy even in cluttered backgrounds and build a model which is comparatively less dependent on light. In future work, we will try to improve accuracy by removing background and performing heavy preprocessing on the images.

References

1. Rao GA, Syamala K, Kishore PVV, Sastry ASCS (2018) Deep convolutional neural networks for sign language recognition. In: 2018 conference on signal processing and communication engineering systems (SPACES), Vijayawada, pp 194–197. https://doi.org/10.1109/SPACES.2018.8316344
2. Taskiran M, Killioglu M, Kahraman N (2018) A real-time system for recognition of american sign language by using deep learning. In: 2018 41st international conference on telecommunications and signal processing (TSP), pp 1–5. https://doi.org/10.1109/TSP.2018.8441304
3. Bantupalli K, Xie Y (2018) American sign language recognition using deep learning and computer vision. In: 2018 IEEE international conference on big data (Big Data), Seattle, WA, USA, pp 4896–4899. https://doi.org/10.1109/BigData.2018.8622141
4. Victoria A, Adejoke O (2018) Conversion of sign language to text and speech using machine learning techniques. 5:58–65
5. Intwala N, Banerjee A, Meenakshi, Gala N (2019) Indian sign language converter using convolutional neural networks. In: 2019 IEEE 5th international conference for convergence in technology (I2CT), Bombay, India, pp 1–5. https://doi.org/10.1109/I2CT45611.2019.9033667
6. Nguyen HBD, Do HN (2019) Deep learning for american sign language finger-spelling recognition system. In: 2019 26th international conference on telecommunications (ICT), Hanoi, Vietnam, pp 314–318. https://doi.org/10.1109/ICT.2019.8798856
7. Khan SA, Joy AD, Asaduzzaman SM, Hossain M (2019) An efficient sign language translator device using convolutional neural network and customized ROI segmentation. In: 2019 2nd

international conference on communication engineering and technology (ICCET), Nagoya, Japan, pp 152–156. https://doi.org/10.1109/ICCET.2019.8726895
8. Kurhekar P, Phadtare J, Sinha S, Shirsat KP (2019) Real time sign language estimation system. In: 2019 3rd international conference on trends in electronics and informatics (ICOEI), Tirunelveli, India, pp 654–658. https://doi.org/10.1109/ICOEI.2019.8862701
9. Tiku K, Maloo J, Ramesh A, Indra R (2020) Real-time conversion of sign language to text and speech. In: 2020 second international conference on inventive research in computing applications (ICIRCA), Coimbatore, India, pp 346–351. https://doi.org/10.1109/ICIRCA48905.2020.9182877
10. Fernandes L, Dalvi P, Junnarkar A, Bansode M (2020) Convolutional neural network based bidirectional sign language translation system. In: 2020 third international conference on smart systems and inventive technology (ICSSIT), Tirunelveli, India, pp 769–775. https://doi.org/10.1109/ICSSIT48917.2020.9214272
11. Beasley WC (1940) Partial deafness and hearing-aid design: I. characteristics of hearing loss in various types of deafness. J Soc Motion Picture Eng 35(7):59–85. https://doi.org/10.5594/J10069.s

Extract It! Product Category Extraction by Transfer Learning

Harsh Raj, Aakansha Gupta, and Rahul Katarya

Abstract The categorization of e-commerce products is salient as its quality directly affects search, recommendations, and related personalized services. Putting the products into the best possible category in a hierarchical manner divided by subcategories is challenging due to the vast range of products creating complexity in product information to select suitable categories. Earlier research proves inefficient with a smaller dataset, so we propose a model to extract the most relevant information from the product description and a pre-trained vocabulary to transform it into subcategories for the prediction of the product category tree. A minute inaccuracy, in this case, can hamper customer satisfaction while searching for the products in the desired categories. To tackle these challenges, we merge the areas of machine learning and deep learning with natural language processing (NLP) to propose a multi-level-based product categorization model. We implemented a selective approach starting with the product name being the node of the tree to recursively form a hierarchical tree structure, searching and extracting subcategories from the product description and a pre-trained vocabulary, which eventually made us predict the most relevant categorization of products. Basically, it is a process of extraction of subcategory levels from the product description and a pre-trained vocabulary built by transfer learning. The proposed model was tested on a Flipkart product dataset containing 20,000 products with several features describing the product. The cosine similarity between the predicted and the given product category tree computed was 0.77, which takes the accuracy to 86%.

Keywords Categorization · Word embeddings · Word extraction · Cosine similarity · Classification · E-commerce

H. Raj (✉) · A. Gupta · R. Katarya
Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

R. Katarya
e-mail: rahulkatarya@dtu.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_9

1 Introduction

With the rapid growth of e-commerce, giant companies like Amazon, eBay, Flipkart, and Rakuten have listed millions of products on their Web sites which are sold by thousands of merchants. From the perspective of a customer, e-commerce means buying products from the comfort of their home To organize products so that customers navigation becomes easy through the searches, building product taxonomy is equally important. “mobile’s accessories>> mobile accessories>> car accessories>> mobile holders>> adroitz mobile holders>> adroitz premium phone socket holder for htc one” is an example for such category trees. With billions of global digital buyers offering countless numbers of products across a wide range of categories, it is required to categorize the products according to their taxonomy by proposing an efficient algorithm to be customer centric and engage into more businesses through the online market. A product taxonomy puts the products into a systematic structure according to the information about products and their use, making it organized while presenting to the users. This standardized arrangement is also valuable for data presentation and storage. Accurate product categorization helps users quickly find the set of products they are looking for. Product categorization in a hierarchical manner helps users get the right results while browsing a category. It also refines the result when a user is doing a keyword search; the product taxonomy helps clarify or disambiguate a search result set based on the categories. Before the advent of machine learning techniques in NLP, merchants and companies were required to manually assign each product to their respective categories, which is a tedious and time-consuming task with error-prone results. Moreover, manual categorization may not be accurate while assigning products to categories as a product can be listed in different ways by different merchants, making this approach inconsistent. Automatic categorization with machines using a consistent algorithm helps to solve these problems. In general, a product taxonomy makes the data collection process more efficient, ensuring Web sites to support features like search refinement, category browse, and specifications while displaying and comparing products which eventually maximizes conversions and profits.

The next question is: How do you improve automatic product categorization models? This question has troubled researchers for a long time, but the best way out we can get is to use machine learning in analyzing and pertaining to the needs of automation. At this scale, even a small increase in the volume of products can result in large calculations to reclassify everything. The investment into more efficient algorithms for product classification is rising every day with the increasing adoption of e-commerce platforms. Newer solutions are trying to determine categories based on:

- Rough product titles and description
- Images of products
- Parsing metadata.

1.1 Our Contribution

To the best of our knowledge, this paper deals with the first kind of product categorization using product titles and its description without explicit training, which can be used even with a smaller dataset or with less information. There are two commonly used approaches; one is the single step classification method which considers all the subcategories as one category, and the other is called the stepwise classifier, which treats different subcategories as different levels. The latter involves a recursive approach to get every subcategory, respectively. Separate model calls are initiated to predict every level of taxonomy in a hierarchical manner. Despite a better accuracy given by the model, it increases the number of times model training has to be performed.

In the past, Kozareva [1] proposed a product categorization approach with derived features such as linear discriminant analysis (LDA) [2], N-gram, Bi-gram, etc. In categorizing the products, the authors experimented with two algorithms: one-against-all (OAA) and error correction tournament (ECT). Recent experiments had been evaluated on a large-scale dataset with as many as 18 million product titles and an extensively larger category list [3]. In our experiment, we have used a transfer learning-based feature extractor to get the most similar words from the product description and a pre-trained vocabulary, starting with product title as the lowest level. We have noticed that the product description contains words that are either in the category tree or words similar to it are present. Hence, our proposed approach of extracting similar words from the description or from the pre-trained vocabulary has proved efficient, giving breakthroughs even with relatively smaller data containing less information as compared to past researches. Besides this, there are problems that hamper fair comparisons with other proposed approaches; most of them do not have publicly available implementations with an evaluation done using private datasets.

2 Related Works

Numerous single-step classifiers have been introduced in the past to use them for product categorization. Yu et al. [4] examined a large number of word-level features using several primitive techniques (e.g., n-grams) and did classification through support vector machine (SVM; Cortes and Vapnik [5]). In the paper [6], Sun et al. use primitive classifiers (e.g., Naive Bayes, k-nearest neighbors) and reduced the errors via crowdsourcing manual labor. Ha et al. [7] and Xia et al. [8] used deep learning to produce word vectors from the product attributes (e.g., product title, product image, merchant ID) and utilized their representation in product categorization.

To improve the categorization results, several stepwise classifiers have also been used. Shen et al. [9] used primitive classifiers such as Naive Bayes and k-nearest neighbors and assigned the product to a leaf node via SVM. Similarly, Das et al. [10] used gradient boosted trees [3] with convolutional neural networks (CNN) for it.

Cevahir and Murakami [11] used deep belief networks [12] and k-nearest neighbors (KNNs) in their approach consisting of two steps.

The paper [13] was among the first to use supervised learning to categorize products into known categories. Their approach simply grouped the product into a particular category with similar products given their informational details (e.g., description, name). They analyzed product information considering information distribution to build features to give input to the classifier. Their results showed significant improvement on results obtained through manual labor. Substantially, they were certainly able to pioneer automation in the categorization of products.

Kozareva [14] introduced a categorization technique with a error correction tournament (ECT) and one-against-all (OAA) OAA reduced the multi-classification problem into multiple binary classification tasks by iteratively classifying each product title for a category and comparing it against all other categories. ECT also reduced the problem to binary classification but employed a single-elimination tournament strategy to compare a set of K players and repeated this process for O(log K) rounds to determine the multi-class label.

In this paper [15], the authors proposed a different paradigm via machine translation. In their model, they translated a product's textual description into a sequence of tokens representing product taxonomy in a root-to-leaf path. They did the experiments on two very large real-world datasets and concluded by showing that their approach was better than traditional classification-based models.

In the study [16], researchers at WalmartLabs compared hierarchical models with flat models for categorizing products. The team deployed deep learning to fetch features from each product to build a product impression. They applied a multi-CNN and multi-LSTM-based approach for this stepwise classification task.

Despite various breakthroughs in the field, there has not been much consideration in improving the models to produce state-of-the-art results with considerably reducing computational complexity even with smaller datasets as compared to the existing approaches.

3 Dataset

The dataset used is a Flipkart e-commerce sample consisting of 20,000 products with their corresponding features, listing the product description and resulting categorization tree in a comprehensive CSV document separating multiple product features. We took a relatively smaller dataset to prove our model's consistency without extensive training.

The dataset includes several other features, but most of them would not contribute to the accuracy of our prediction but could increase the complexity and inefficiency of the task. We removed duplicate product listings leaving us with Flipkart product titles and descriptions which are related to their multi-level category labels. As inferred from the data of most e-commerce sites, the product description and titles contain most of the words present in the category tree of the respective products.

Subsequently, we took product name and description as the features in our approach containing the information required for the prediction of the taxonomy tree. During training and evaluating our approach, we had split the data into a training set of 16,000 samples and a test set consisting of 4000 samples. We predicted the product taxonomy up to five hierarchical levels.

4 Methodology

We explored various approaches for the tasks and looked more closely at the correlation between the features and the target. We found that the correct extraction of words from the features is the key in predicting the product category.

This section deals with the methodology of our model, starting with preprocessing and selecting the important features; our proposed model first preprocesses the textual features by removing punctuations and numbers and converting words to their lowercase. After conducting experiments with other preprocessing techniques such as stemming and lemmatization, we saw a relative decline in the final accuracy of our model; it might be because the predicted product categories were losing their local context.

4.1 Preprocessing

In our proposed approach, we applied preprocessing steps on product title and description. This is done to extract most of the information about the product and to increase the materiality of the texts describing it in order to predict the taxonomy tree, which may be lost without this step.

Number removal: On noticing the features describing the products, we realized that numbers had very little relevance to a given product category, so we decided to remove them, considering that the discrete distribution of numbers would only adversely affect the given categorization results. We believe that it is not possible to fully capture the information the numbers try to pursue for a given category. For example, 1–9, 3, 6, 0.3, we expect the fact that for certain products, specific numbers may show up more relevance, whose inconsideration will eventually affect the categorization. But to get the vectors of the word, number removal is a necessary step.

Removing punctuation: It was done as a result of manual examination of the data through which we noticed that humans have various options for choosing words and phrases, as a word may have multiple accepted forms. Also, there are punctuations in between the phrases, including commas, parenthesis, question mark, exclamation points, and others. Also, in most cases, concatenated words are used in different senses. Ex: won't, wont, 2D, 2-D, good., good? Removing punctuation resulted in

being one of the most effective feature normalization techniques used to get a clearer context of information through words.

Lowercasing: According to the standards of the English language, sentences should have the first word capitalized. Humans, while writing, often capitalize some words ambiguously, depending on their intent, interpretation, or choice of capitalizing acronyms. Precisely, this cannot always be a useful scaling technique, as it can be contradictory, especially when dealing with product names vs. generic objects at the same time—e.g., Home, home, CD, cd, Windows, windows. As observed, lowercasing does not help much, most likely because of the shift in meaning caused by lowercasing occurrences such as "Schools are good" to "Schools are good" are offset by "Windows XP" to "windows xp", which have very different meanings.

4.2 Feature Selection

As our dataset contains a comprehensive set of features to describe products and their usage, not all of them will be useful to be applied for classification or differentiation between categories. We only considered those features which contain most of the information about the product, so we decided to use feature selection to achieve this. We first looked at the distribution of words of subcategories of the product category tree in every feature using the Jaccard index [17] and prioritized the features which could be considered as important for the task. The value of the Jaccard index was maximum for the following features: description (0.31) and product name (0.42). Therefore, we conducted our experiments using these two features, i.e., description and product name.

Using the product description as the most informative feature increased the accuracy of our model. When we also use the least informative features, the accuracy drastically reduces to get halved. This might be because the mutual information would not be capturing the least informative feature due to the lack of distribution of information of categories in them.

The product name feature refers to the name of the products, which is crucial in determining the taxonomy tree as it will fill the lowest level in the product taxonomy tree. Product name will be used in the extraction of other subcategories as it will initiate the recursive extraction process from the product description and a pre-trained vocabulary.

The target feature for our dataset is the product category tree. It is divided into hierarchical sublevels (subcategories) with the use of " >> ", punctuation which got removed in the preprocessing step. The uniqueness of our proposed model is that it predicts the taxonomy tree just by modifying the extraction of words from the description and a pre-trained vocabulary with finer accuracy even without explicit training.

4.3 Predicting the Taxonomy Tree

This section will introduce our proposed method for extracting the taxonomy tree to get a hierarchical level-based categorization of products, starting with the product name as the first level(subcategory) and recursively searching its most similar(suitable) words from the product description and a pre-trained vocabulary.

4.3.1 Feature Representation

There is a need to get the vectors of words to be used in any machine learning model; in our paper, we used Word2Vec [18], GloVe [19], FastText [20], Paragram [21] word embeddings for the purpose. The method used transfer learning where the pre-trained embedding model was used to get the vectors of texts without explicitly training it for the task.

Word2Vec: Word2Vec [18] being a very popular pre-trained word embedding model was developed by Google. They have also been applied to various tasks such as recommendation systems, knowledge extraction and discovery, and different problems related to text analysis. Word2Vec model architecture contains a feed-forward neural network with one hidden layer. Therefore, it is also known to have a shallow neural network architecture. For the purpose of vector conversion of words, we pre-trained the Word2Vec [18] model, producing a sequence of 300-dimensional embeddings for various words and phrases.

GloVe: GloVe [19] is an unsupervised learning algorithm used for word vectorization. While training, a co-occurrence matrix is created which represents the linear substructures of the word in its vector space w.r.t. the occurrence of other words. The objective for training is to minimize the difference in dot products of word vectors and the logarithm of the probability of co-occurrence of those words. Subsequently, the ratio in a logarithm is same as the difference of the logarithms, so this objective is ultimately the ratios of co-occurrence probabilities with vector differences in a space of vectors. Because of this, the word vectors perform extraordinarily on tasks related to word analogy and correlation. For the purpose of word vectorization, we pre-trained the GloVe [19] model producing a sequence of 300-dimensional embeddings for words and phrases.

FastText: Abovementioned methods ignore the morphology of words, while creating vectors for each word in the vocabulary. The main limitation comes with syntactically rich languages as these kinds of languages have larger vocabularies and a relatively complex syntax, yielding lower-quality representations. Bojanowski et al. [20] in 2016 introduced a model on this principle of Skip-gram called FastText which perceives each word as a bag of n-gram characters. It learns the embeddings for each word by taking the sum of its corresponding n-gram embeddings. For text vectorization, pre-trained FastText [20] model was used which produced a sequence of 300-dimensional embeddings for each word and phrase.

Paragram: Proposed by Wieting et al. [21] is a model for word embedding for learning with pairs of paraphrase from PPDB [22] database. Precisely, this method encodes phrases to a vector space by minimizing the difference in the cosine similarity in the space with the scores of pairs of paraphrase phrases. For text vectorization, pre-trained Paragram [21] model was used which produced a sequence of 300-dimensional embeddings for each word and phrase.

The pre-trained model used in word vectorization is able to produce a large vocabulary, which is used in refining our model's accuracy in predicting the product category tree.

4.3.2 Weighted Word Embeddings

This section deals with the approach devised to find text embeddings for product names. Most of the product names are not in the form of words but are represented as texts containing words, and to treat the product name as the first level in the product category tree, we have to represent it as a single vector.

To calculate the text embeddings, our model takes the term frequency-inverse document frequency (TF-IDF) [23] mean of word embedding. For each product name, we calculated a weighted mean of the word vector obtained using the Word2Vec [18], GloVe [19], FastText [20], and Paragram [21], where the weights are the TF-IDF value of the word.

4.3.3 Evaluation Metric

We have used cosine similarity to measure the similarity to recursively produce the product taxonomy tree by extracting the most similar words. Cosine similarity [24] is chosen as the measure as it gives accuracy based on the contextual as well as semantic relations between the input vectors. It measures the similarity between two vectors by taking the cosine of the angle to determine the degree of closeness. For NLP tasks, it is mainly used to measure similarity in textual data. The similarity score ranges from -1 to 1, with the latter being a perfect score.

To select a word (from the description or the pre-trained vocabulary) as the next subcategory in the taxonomy tree, the similarity score between the vectors of the previously predicted subcategory and the product description should be greater than a predetermined threshold value. We also marked the words so that no subcategory gets repeated. Searching from the pre-trained vocabulary over the product description is introduced to tackle the unavailability of similar words.

If our algorithm is not able to search for any word in the text (description) whose similarity score tends to be greater than the predetermined threshold value, then we took the most similar word from the vocabulary, which is not equal to any of the preceding subcategories as the next level in the taxonomy tree.

On exploring, we found that most of the product category trees were having five subcategories on average. So we repeated the above steps 5 times to get the product

Fig. 1 Comparison of predicted against given product category trees

taxonomy consisting of five hierarchical levels, which predicted efficient results with relatively fast execution and low complexity (Fig. 1).

5 Results

This section describes the results for the sets of experiments conducted. As the product taxonomy tree should be evaluated both on semantic and syntactic similarity, the final evaluation was done on both cosine similarity [24] and F1-score. The model training is done to predict taxonomy tree which shall minimize the loss of product names, so a combination of both the scores gives a better scheme for evaluating the method.

The comparison in accuracy for different feature representations techniques (Word2Vec [18], GloVe [19], FastText [20], Paragram [21]) was done based on the cosine similarity [24] and the F1-score, and the results are shown in Table 1.

While evaluating, we noticed that the precision is always higher than recall which means that the result does not contain out of context words.

Table 1 Results comparing paragram, Word2Vec, GloVe, and FastText embedding techniques

Feature representation	Cosine similarity	Precision	Recall	F1-Score
Paragram	0.78	0.77	0.71	0.74
Word2Vec	0.83	0.79	0.71	0.75
GloVe	0.86	0.85	0.74	0.79
FastText	0.86	0.88	0.79	0.83

6 Conclusion

In the paper, we have presented a product categorization model on the basis of extraction of sub categorical levels from the product description and a pre-trained vocabulary of about 3 million words and phrases with their vectors, which significantly improved the accuracy correlating the predicted and given category tree. We classified products into a five-level taxonomy. Our experiments showed that GloVe [19] and FastText [20] embeddings-based extraction of product taxonomy lead to the best performance reaching 0.86 and 0.82 in cosine similarity [24] and F1-score, respectively. We also manually examined the produced categorization outputs and found that often our predicted results are more specific and fine-grained in comparison with those provided manually.

As the word vectors created were pre-trained on general text corpus extracted from the web, so in some cases they were not able to grasp the context of the words in the product names specific to the e-commerce industry. Pre-training these word embeddings on task-specific text corpus might help in improving the vectorizarion of given texts and thus the model performance.

References

1. Campbell JC, Hindle A, Stroulia E (2015) Latent dirichlet allocation: extracting topics from software engineering data. Art Sci Anal Softw Data 3:139–159. https://doi.org/10.1016/B978-0-12-411519-4.00006-9
2. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
3. Friedman JH (2000) Greedy function approximation: a gradient boosting machine. Annals of Stat 29:1189–1232
4. Yu H, Ho C, Arunachalam P, Somaiya M, Lin C (2012) Product title classification versus text classification. Csie.Ntu.Edu.Tw, pp 1–25. [Online] Available http://www.csie.ntu.edu.tw/~cjlin/papers/title.pdf
5. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
6. Sun C, Rampalli N, Yang F, Doan AH (2014) Chimera: large-scale classification using machine learning, rules, and crowdsourcing. Proc VLDB Endow 7(13):1529–1540. https://doi.org/10.14778/2733004.2733024
7. Ha J-W, Pyo H, Kim J (2016) Large-scale item categorization in e-commerce using multiple recurrent neural networks. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp 107–115
8. Xia Y, Levine A, Das P, Di Fabbrizio G, Shinzato K, Datta A (2017) Large-scale categorization of japanese product titles using neural attention models. In: Proceedings of the 15th conference of the european chapter of the association for computational linguistics: Volume 2, Short Papers. Association for Computational Linguistics
9. Shen D, Ruvini J-D, Sarwar B (2012) Large-scale item categorization for e-commerce. In: Proceedings of the 21st ACM international conference on information and knowledge management, CIKM '12, New York, NY, USA. ACM, pp 595–604
10. Das P, Xia Y, Levine A, Di Fabbrizio G, Datta A (2016) Large-scale taxonomy categorization for noisy product listings. In: Big Data (Big Data), 2016 IEEE international conference on, IEEE, pp 3885–3894

11. Cevahir A, Murakami K (2016) Large-scale multi-class and hierarchical product categorization for an e-commerce giant. In: Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers, pp 525–535
12. Hinton GE, Osindero S, Teh Y-W (2006) A fast learning algorithm for deep belief nets. Neural Comput 18(7):1527–1554
13. Chen J, Warren D (2013) Cost-sensitive learning for large-scale hierarchic al classification. Int Conf Inf Knowl Manag Proc 1351–1360. https://doi.org/10.1145/2505515.2505582
14. Kozareva Z (2015) Everyone likes shopping! multi-class product categorization for e-Commerce. In: NAACL HLT 2015–2015 conference of the north american chapter of the association for computational linguistics: human language technologies proceedings conference, pp 1329–1333. https://doi.org/10.3115/v1/n15-1147
15. Li MY, Kok S, Tan L (2018) Don't classify, translate: multi-level e-commerce product categorization via machine translation. arXiv, pp 1–15
16. Krishnan A, Amarthaluri A (2019) Large scale product categorization using structured and unstructured attributes. arXiv, 2019
17. https://en.wikipedia.org/wiki/Jaccard_index
18. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J (2013) Distributed representations of words and phrases and their compositionality. Adv Neural Inf Process Syst 1–9
19. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation Proc. EMNLP 12:1532–1543
20. Bojanowski P, Grave E, Joulin A, Mikolov T (2016) Enriching word vectors with subword information. arXiv:1607.04606.
21. Wieting J, Bansal M, Gimpel K, Livescu K, Roth D (2015) From paraphrase database to compositional paraphrase model and back trans. ACL 3:345–358
22. Ganitkevitch J, Van Durme B, Callison-Burch C (2013) PPDB: the paraphrase database. In: Proceedings of HLT-NAACL, pp 758–764
23. Arora S, Liang Y, Ma T (2016) A simple but though baseline for sentence embeddings. Iclr 15:416–424
24. Faisal R, Kitasuka T, Aritsugi M (2012) Semantic cosine similarity. Semantic Scholar 2(4):4–5

Readability of Websites for the Novel Coronavirus Disease (COVID-19): A Case Study of India

Diksha Dani and Gaurav Agrawal

Abstract The World Wide Web has become one of the most important resources for the large part of the population in searching healthcare information. The available COVID-19 information on Websites should be reliable and accurate in order to spread awareness. This has become all the more important in times like COVID-19, pandemic situation when it is necessary to create awareness about the disease to large population in a country like India and also provide updated information and status of the situation. However, readability poses barrier for many people. In this paper, readability of online information related to the novel coronavirus disease (COVID-19) is evaluated using online readability tool. It was found that the readability of content addressing COVID-19 is satisfactory. But the absence of multilanguage option in majority of Websites under study forms another barrier in information dissemination.

Keywords COVID-19 · Readability · WebFX tool

1 Introduction

India is a country with the second largest population in the world. The COVID-19 disease that has brought the world to its knees is highly contagious, and social distancing has been suggested as one of the measures to prevent its spread. Owing to India's large population, it is difficult to maintain social distancing; consequently, the disease has a higher probability of spreading faster. Also, due to social media platforms, false information and rumors may spread fast. So, communicating the correct updated information about preventive measures to the maximum possible people is a massive task. The Internet has proved to be quite a boon in such situations. But, there is another issue with this task. India is a land of diverse languages and dialects, and the majority of the population lives in rural areas. Therefore, it is equally important that the information available online on highly communicable disease like COVID-19 is to be conveyed in such a manner that the majority of the population

D. Dani (✉) · G. Agrawal
Inderprastha Engineering College, Ghaziabad, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_10

understands. Since, due to the spread of digital information deep even in remote areas, the Internet has become the primary source of information regarding awareness and treatment of patients, it is essential to evaluate the readability levels of the information presented by the online Websites. This study evaluates the readability of information available online related to the COVID-19 pandemic in Indian Websites.

2 Related Work

There are many studies on the readability of healthcare material available online. Grose et al. [1] evaluated the readability and quality of online information on septoplasty. The authors concluded that most online education materials on septoplasty have significant quality and reliability issues and severely high reading levels. Likewise, in another study [2], online education materials on congestive heart failure for patients were evaluated for readability. Out of 70 Websites, only 5 met the recommended sixth-grade readability level.

However, there are not many papers explicitly about the COVID-19 pandemic on the Internet. The authors of the paper [3] assessed the readability of information available over the Internet about the novel coronavirus disease (COVID-19) to determine compliance with reading level standards. They found that online content about COVID-19 is complicated to read and understand, as advocated by the American Medical Association (AMA) and the United States Department of Health and Human Services. Kruse et al. conducted another study linked to COVID-19 reading [4]. The researchers assessed the content of the Website, readability of the text present on site, and quality of Web-based PEMs from U.S. academic medical centers on COVID-19. They discovered that the mean readability was greater than the suggested sixth-grade reading level set by the National Institute of Health and the U.S. Department of Health and Human Services. Worral et al. [5] evaluated the readability of COVID-19-related online information of four countries, namely: Canada, The United Kingdom, Ireland, and the United States, and compared the readability of Website source origin and geographical origin. Only 17.2% of Webpages were found to be widely accessible. Vishala and Dexter [6] assessed the readability of COVID-19-related online content given by the government and public health agencies and departments. They discovered that all 18 Websites under study were easy to understand and had the reading level of grades 6 through 8.

3 Methodology

In this study, the readability of COVID-19-related Websites in India was evaluated. The Website or Webpage was searched by Google using the keywords: like “COVID-19-related Websites in India” or “COVID-19 information in India” or “COVID-19 resources in India.” A total of 38 Websites belonging to various cities, states, and

centers were evaluated. The choice was toward selecting the public health Websites of the COVID-19 pandemic as of mid-May 2021. The evaluation was conducted on the English language of the homepage of every Website. If the Website was in any other language, it was translated to English before applying the test.

COVID-19 is a pandemic that has spread over both rural and urban areas. Therefore, it was required that the information be disseminated quickly to all the people, so it was essential to evaluate the reading ease of the text written on the Website. Readability is a metric for determining how easy or difficult a text is to read and comprehend. Since human evaluation is prone to errors and inconsistencies, so to ensure consistency and avoid errors caused by humanly limitations, the readability scores were computed using an online readability tool. There are various online tools available for computing readability. The online readability calculator [7] was used to calculate the readability of Websites to avoid human evaluation errors. The tool takes URL as an input, or plain text can also be copied and pasted into the tool to evaluate the readability score. In this paper, the URL of the Website under study is given as input to the online tool.

The various readability parameters, along with the definitions and formula, are taken from the WebFX tool as follows:

Flesch Kincaid Reading Ease (FKRE) test measures how easy or difficult the text is to understand. The measure is based on a score between zero and a hundred. A lower score indicates that the text is difficult to understand, while a higher score indicates the ease of understanding the text. The measure uses formula 1 to calculate the reading score.

$$\text{FKRE score} = 206.835 - 1.015 \times \left(\frac{\text{Words}}{\text{Sentences}}\right) - 84.6 \times \left(\frac{\text{Syllable}}{\text{Words}}\right) \quad (1)$$

Flesch Kincaid Grade Level (FKGL) test: The Flesch Kincaid Grade Level is used to assess a text's approximate reading grade level. The score obtained indicates the grade level. For example, if the grade number is six, an average student in the sixth can read and understand the content with ease. It is calculated using the formula 2:

$$\text{FKGL Score} = 0.39 \times \left(\frac{\text{Total Words}}{\text{Total Sentences}}\right) + 1.8 \times \left(\frac{\text{Total Syllables}}{\text{Total Words}}\right) - 15.59 \quad (2)$$

Gunning Fog (G.F.) score: The Gunning Fog Index is a metric that determines the number of years of education required to comprehend a passage of text on first reading. The Gunning Fog algorithm yields a grade level score ranging from 0 to 20. The grade level is determined by the score received. A score of 6 indicates that the text is appropriate for a 6th grader. It is a weighted average of how many words per phrase and how many lengthy words per word are there in the sentence. It is computed using formula 3:

$$\text{Gunning Fog Score} = 0.4 \times \left[\left(\frac{\text{Total words}}{\text{Total Sentences}} \right) + \left(\frac{\text{Complex Words}}{\text{Total Words}} \right) \right] \quad (3)$$

Coleman Liau Index (CLI) is another index of readability measured using the average number of characters and sentences per 100 words. The grade level is the floor and ceiling values of the score obtained. For example, if the score is 9.3, it is understood by 9th and 10th students. It is mainly used to assess educational text material but can be used for other text too. It is calculated using formula 4:

$$\text{Coleman Liau Index(CLI)} = 5.89 \times (\frac{\text{characters}}{\text{words}}) - 0.3 * (\frac{\text{sentences}}{\text{words}}) - 15.8 \quad (4)$$

Automated Readability Index (ARI): The automated readability index is a readability test to measure how easy it is to comprehend the text. A score of 1 means it can be understood by Kindergarten grade students of age 5–6 years old while a score. Similarly, a score of 12 means 12th-grade students can understand it, while the score of 18–24 can easily be understandable by the college student. The professor-level people can comprehend a score greater than 24. It is calculated using formula 5:

$$\text{Automated Readability Index (ARI)} = 4.71 \times \left(\frac{\text{characters}}{\text{words}} \right) + 0.5 \times \left(\frac{\text{words}}{\text{sentences}} \right) - 21.43 \quad (5)$$

SMOG Index: SMOG, a readability index, stands for "Simple Measure of Gobbledygook," which estimates how many years of education the average person needs to understand a text. It is calculated using the formula 6.

$$\text{SMOG Index} = 1.0430 \times (\sqrt{30 \times \left(\frac{\text{complex words}}{\text{sentences}} \right)}) + 3.1291 \quad (6)$$

4 Results

The results indicated that 52% of the Websites fell under the easy to read category, while 48% of Websites had a reading level of high school and above. The mean readability score for FKRE, FKGL, G.F., SMOG, CLI, ARI was 58.97, 7.04, 6.9, 7.25, 8.9, 2.9, respectively. Table 1 shows the statistical results of readability scores of the Website under study. Figure 1 shows the score of the various readability indices applied on the Website under study. The mean FKRE score of 58.97 indicates a "fairly difficult to read" score and is readable by 10^{th} -11^{th}-grade students. The highest FKRE score was reported was 101 by the Website https://covidinfo.rajasthan.gov.in. A score

Table1 Descriptive statistics of the readability indices of the websites

	N	Minimum	Maximum	Mean	Std. deviation
Flesch kincaid reading ease	38	6.50	101.50	58.97	21.37
Flesch kincaid grade level	38	0.70	31.60	7.04	4.83
Gunning fog score	38	2.80	31.80	6.91	4.85
SMOG index	38	3.40	18.80	7.26	2.68
Coleman liau index	38	−9.50	18.90	8.97	5.87
Automated readability index	38	−13.10	27.60	2.99	6.29

of 100 or more means the text is very easy to read and can be understood by 11 year old.

The public health professionals recommend a 6th-grade reading level to be a universal reading level [8]. Table 2 shows the result of FKRE score on Websites under study. The lowest FKRE score reported was 6.5 by https://covid19.assam.gov.in/ Website, which is interpreted as very difficult to read and can be understood only by college graduates. The mean readability was found to be of 6th and 7th g**rade.**

But another issue in India regarding online content dissemination is the information should be presented in native language. Table 3 shows the result of presence of multi-lingual option on Websi**te.**

In one of the studies, it was found that Website languages may severely limit Website access by users who are not provided with information in their native language [9]**.**

In a country like India, where there are 33 different languages and 2000 dialects [10], out of 38 Websites that were evaluated, 19 Websites were only in English while 15 had multilingual support, 3 Websites were in Hindi as well as in English while one Website was in English and a regional language. Figure 2 shows the result of readability indices of Websites under study.

5 Threat to Validity

Many of these tests are based on English language text, and the result may not be significant if the text is in a language other than English. Also, the scores obtained indicate the U.S. school level required by a person to understand the text. How much valid they are in the Indian context is another question to ponder. Also, different readability index gives different score and grade for the same text. Additionally, readability does not consider the meaning and context of the text. Low literacy levels in India are another barrier to online health knowledge.

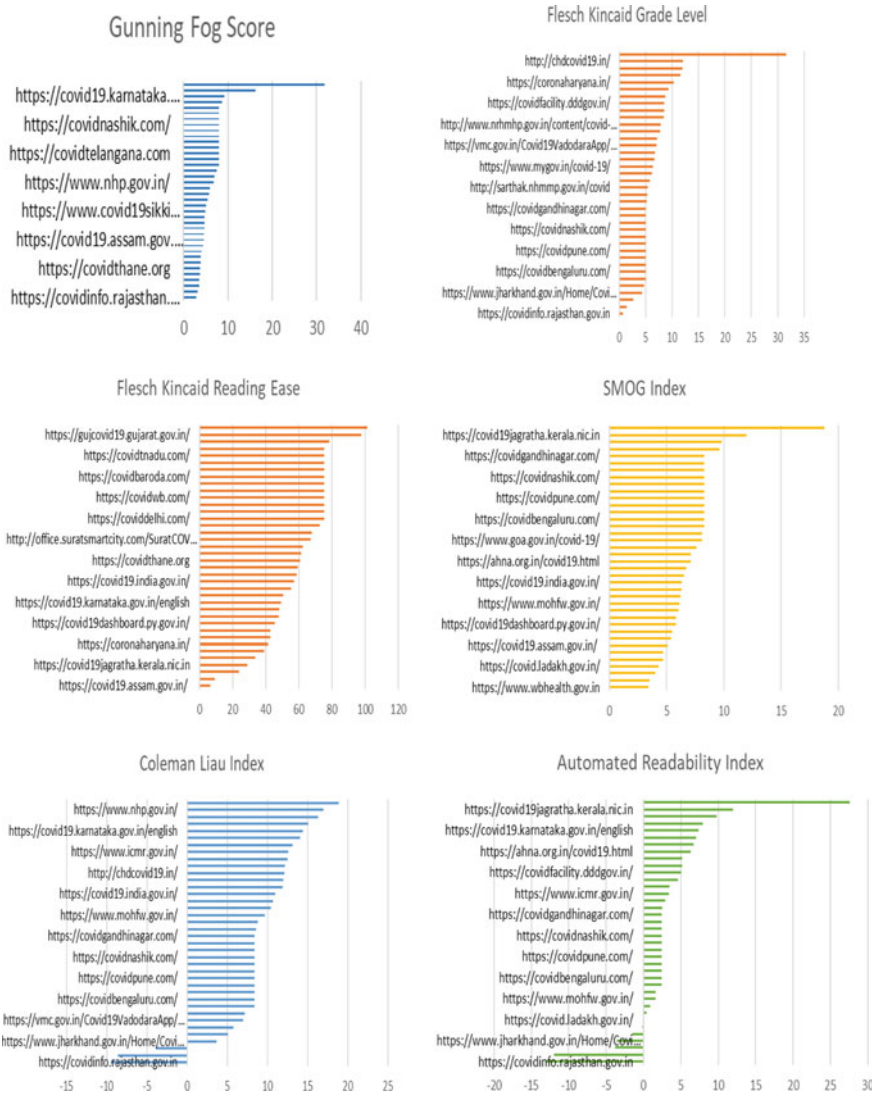


Fig. 1 Bar chart showing representation of readability indices

6 Conclusion

A large population of India uses the Internet for healthcare information. This is due to the digital revolution, which has made the Internet accessible even in remote areas. There is a vast amount of COVID-19 content available online. In the study, it was found that the readability of content addressing COVID-19 is satisfactory. But language forms another barrier in information dissemination. Research is needed to

Table 2 Number of websites falling in different categories according to FKRE

Score	School level	Interpretation	No. of websites
90.0–100.0	5th grade	Very easy to read	2
80.0–90.0	6th grade	Easy to read	0
70.0–80.0	7th grade	Fairly easy to read	13
60.0–70.0	8th and 9th grade	Plain english	5
50.0–60.0	10th to 12th grade	Fairly difficult to read	5
30.0–50.0	College	Difficult to read	9
0.0–30.0	College graduate	Very difficult to read	4

Table 3 Number of websites in different language

Total sites	38
Only English	19
Hindi/English	3
Multilingual	15
Regional	1

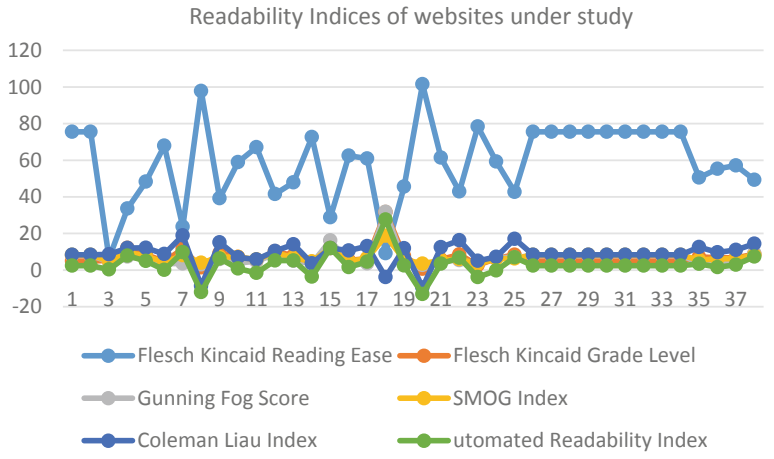


Fig. 2 Readability indices of websites

make the readability measures adapt to multilingual pages focusing on context and its meaning in the language.

The pandemic nature of COVID-19 requires adequate and effective communication and dissemination of reliable information from the healthcare people to the general public. This can be done by creating online material written at the recommended reading level and comprehensible to help the public take informed decisions about risk mitigation measures.

References

1. Grose EM, Holmes CP, Aravinthan KA, WV ,Lee JM (2021) Readability and quality assessment of internet-based patient education materials related to nasal septoplasty. J Otolaryngol—Head Neck Surg 50(16)
2. Kher A, Johnson S, Griffith R (2017) Readability assessment of online patient education material on congestive heart failure. Adv Prev Med 2017:9780317. https://doi.org/10.1155/2017/9780317
3. Szmuda T, Özdemir C, Ali S, Singh A, Syed MT, Słoniewski P (2020) Readability of online patient education material for the novel coronavirus disease (COVID-19): a cross-sectional health literacy study. Public Health 185:21–25
4. Jessica K, Paloma T, Tayler BB, Erica JT, Charlesnika TE, William AG, Emily SM, Elizabeth MSL(2021) Readability, content, and quality of COVID-19 patient education materials from academic medical centers in the United States. American J Infect Control 49(6):690–693. ISSN 0196–6553. https://doi.org/10.1016/j.ajic.2020.11.023
5. Worrall AP, Connolly MJ (2020) O'Neill A (2020) Readability of online COVID-19 health information: a comparison between four English speaking countries. BMC Public Health 20:1635. https://doi.org/10.1186/s12889-020-09710-5
6. Mishra V, Dexter JP (2020) Comparison of readability of official public health information about COVID-19 on websites of international agencies and the governments of 15 Countries. JAMA Netw Open. 3(8):e2018033.https://doi.org/10.1001/jamanetworkopen.2020.18033
7. Webpagefx. https://www.webfx.com/tools/read-able/. Accessed May'2021
8. McKenzie JF, Neiger B, Thackeray R (2017) Planning, implementing, and evaluating health promotion programs: a primer, 7th edn. Pearson, New York
9. Kralisch A, Berendt B (2005) Language-sensitive search behaviour and the role of domain knowledge. New Rev Hypermedia Multimedia 11:221–246
10. Facts about India-Indian languages. http://www.webindia123.com/india/people/language.htm

ReMeDi: Secure Electronic Health Records Sharing Using Blockchain

Pooja Tripathi, Rohan Yadav, Shruti Pancholi, Vivek Singh, and Nishtha Singh

Abstract The following paper introduces a web application using blockchain technology that put forward an effortless solution for sharing sensitive medical records in a secure way. With the fast pace life involving work and degrading environmental conditions, we are subjected to various health issues. These health issues generate a lot of health records. Health records store a lot of critical information that is private to individuals. Medical records are currently handled manually by individuals which have security and efficiency-related issues. Thus, the web application provides instantaneous tracking of medical records which improves security, increases efficiency, and reduces overall cost. It will help infrequent distribution of information among health care providers, insurance companies, pharmacists, researchers, patient families leading to ease in data sharing, improving decision-making for medical care, and maintaining accessibility and authenticity. This paper highlights this application as it works on two different levels: patient and the doctor levels. While patients can use it for uploading their records, granting access to the doctors, and viewing their records, doctors can use it to view the records of the patients who have given access to them.

Keywords Accessibility · Blockchain · Confidentiality · Doctors · Electronic health records · Patients · Privacy · Security · Web application

P. Tripathi (✉) · R. Yadav · S. Pancholi · V. Singh · N. Singh
Department of Information Technology, Inderprastha Engineering College, Ghaziabad, India
e-mail: pooja.tripathi@ipec.org.in

R. Yadav
e-mail: Rohan.yadav@ipec.org.in

S. Pancholi
e-mail: shruti.pancholi@ipec.org.in

V. Singh
e-mail: vivek.singh@ipec.org.in

N. Singh
e-mail: Nishtha.singh@ipec.org.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_11

1 Introduction

Blockchain emerged as distributed technology after Bitcoin white paper was published [1]. Blockchain is decentralized and has all the transactions in a sequential manner which can be referred to by the users at any point in time without any hassle of a central governing body [2]. This technology assures security of knowledge, regulation of susceptible data between patients and medical field. Various benefits associated with this technology are it is immutable, maintains integrity, is hassle-free and transparent [3]. All these features were possible because of cryptography concepts and various protocols which include digital signature and hashing.

Blockchain has its foot in every possible field including health care. In health care for managing records, as is simple and safe due to the different features of blockchain-like transparency and data integrity [4–6]. It makes patients the sole owner of their data and has full control of their access rights [7]. This technology can also be used to perform various medical and clinical researches as this process breaches the integrity and privacy of patient's medical information [8]. And in such cases, blockchain can work as a solution for all such problems [9].

Health care researchers are working hard to dispose of all such troubles associated with the traditional technology. Blockchain has taken the entire industry by storm health care will be further progressed if accompanied by machine learning and artificial intelligence [10, 11].

It can further be used for fraud detection in the supply chain of medicinal drugs. It is a crucial issue to protect these drugs from any fraud as most people depend on them for life. Any settlement made on this can affect the health. These chains are prone to such fraudulent activities as it involves a lot of people [12, 13].

Blockchain terminates most of the problems of traditional systems that can provide a safe platform; every record in blockchain can be authenticated via smart contracts as and updating blockchain is not easy.

In this fast-paced life, health has been demeriting rapidly, and to get rid of them, it requires various tests and medical records. Always when we visit the doctor, we always have to put forward the reports which include medical records, MRIs, X-rays, previous prescriptions. It is one of the biggest tasks to keep all such records together.

Health records are very sensitive information, and it is required to be divided among insurance companies, doctors, pharmacists, and researchers; this also is very challenging as patients' history keeps on updating. Someone who has a serious disease such as cancer or HIV, they have a history of treatment rehabilitation and post-treatment [14, 15]. Having a complete history is very important for correct treatment, for instance, knowing the dose of chemotherapy for further treatment is very important as the patient can visit several medical facilities for consultation, or they can be transferred from one hospital to another. Traditionally, if a patient wants to share their clinical information for research purposes or even if they want to transfer them from one hospital to another, they have to sign a declaration stating that what data they are agreeing to share with the recipient.

Collection of data for research purposes also requires patient's consent unless it is anonymous. Relying on a single authority that stores data can have a single point of failure and can collapse the backbone of the entire framework. Existing traditional systems have various loopholes such as integrity, privacy, security, and single point of failure as those were centralized. Our paper aims to suggest a framework that proposes a decentralized system that stores patients' medical records, and they can even give access to records to particular doctors or researchers. We even tried to solve the problem of privacy for patients. We even solved the problem of blockchain scalability by using IPFS technology which is an off-chaining method, for storing patient's data.

The paper is divided into different sections. Section 2 summarizes literature review. Section 3 narrates preliminaries like technologies used. Sect. 4 shows the design of the framework and methodology. The last section denotes conclusions and future works.

2 Literature Review

Several such applications have been developed keeping in mind the secure sharing of medical resources. All the applications that used technology to their pre-eminent clarify this problem of secure sharing of medical records. A company in Estonia, Gaurdtime also uses the framework based on blockchain to secure the records of convalescent. An application MedRec [16] built on Ethereum with its smart contracts is to create a brilliant presentation of medical records within nodes of networks. It was an outline by Massachusetts Institute of Technology Research Laboratory in collaboration with BIDMC. It explains the distributed approach for managing access rights authorization and exchange of medical records among health care systems. Blockchain gives the patients complete rights over their data, and the permissions are with patients only. MedRec project has worked fine for medical data, and now, the developers are working to add more scope.

The author presents the Medicalchain framework which eases exchange of health history across medical institutions [17]. Another project Healthcoin directs to construct an EMR system globally. Some other players in this field are Patientory, Factom, IBM Watson, Bowhead, Nuco, SimpleVital, HealthCombix, and QBRIC.

One other framework Ancile [18] also utilizes smart contracts in Ethereum-based blockchain for increased access control and modification of data; it even uses cryptographic techniques for improved security. Healthchain [19] is a framework based on Hyperledger that works on the preservation of privacy. Zhang and Lin moved toward seclusion and put forward seclusion maintaining scheme which used private blockchain to store health records [20]. This framework focuses on asymmetric encryption and also implements procedures for testing of performance to ensure system availability. Another arrangement was suggested via Google platform that offers Google app script a programming facility that helps to integrate wallets and data repository.

MedBlock [21], BlockHIE [22], FHIRChain [23], MeDShare [24], HBasechainDB [25] are other frameworks that are based on blockchain. The authors proposed a platform that is seclusion preserving MediBchain engaged cryptographic functions to identify patients' history in a blockchain system [26]. Yeu proposed Healthcare Data Gateway (HDG) [27] an application that allows patients to control and choose how to share their data by preserving seclusion.

A similar architecture was proposed for sharing and managing medical data on diabetes patients using blockchain contracts using multi-signature to improve data privacy and access control [28]. Griggs, he demonstrated how smart contracts on the Ethereum platform can support an automated secured environment that monitors patients in real time [29], while Liang presented a Hyperledger-based framework [30] for data collection and sharing among different stakeholders in mobile healthcare interface [31]. The proposed framework is flexible to overcome technical issues. In the future after creating a reliable heath network, we could open other frontiers to health management by integrating pharmacists, insurance companies, and researchers.

After surveying various papers and reading about all the different applications, we conclude that in all the mentioned applications, patients are not even considered the owners of their medical records, which is very personal and delicate. Since the patient is the sole owner of the data, they should only have permission to give access to anyone. Mainly, our paper focuses on making patients the sole owners of their medical records. All the powers regarding the data are in their hands. ReMeDi uses the access control mechanism through which patients can give access to their medical records, and later, the access can easily be revoked, and this makes ReMeDi different from already existing applications. ReMeDi incorporates a cryptocurrency wallet, which will directly consolidate the transaction cost deductions, making it unique, highly scalable, and efficient.

3 Methodology Used for Framework

Designing of the Framework

ReMeDi does not store the patient's records directly into the blockchain, but it stores the patient's records on IPFS, and the cryptographic hash is stored into the blockchain. Thus, for storing records, the off-chaining technology is used that is accomplished through IPFS. It is the hybrid approach because files are not stored on the blockchain directly, but the metadata of the record is stored on the blockchain that can be used to access the records. The overall working of the framework is defined in Fig. 1.

Implementation of the Framework

The framework is divided into two main components:

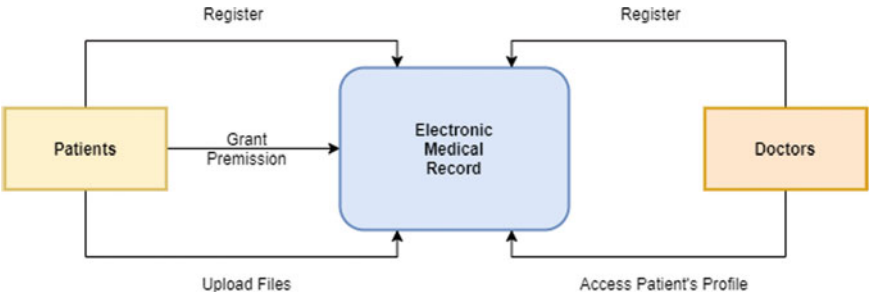


Fig. 1 Working of framework

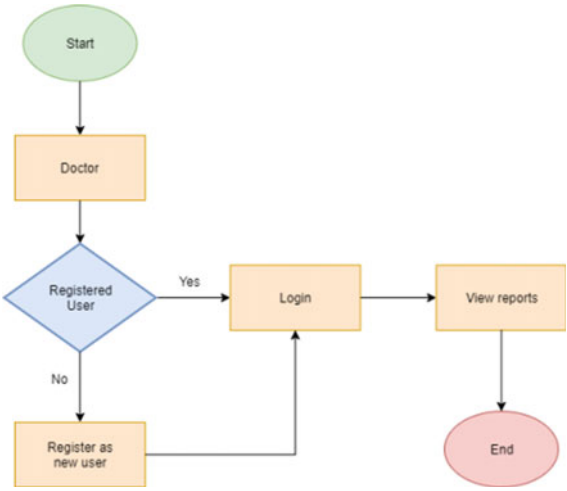
Doctor Component

For his part, the doctors can register themselves and then sign into their personal dashboards. Only once the particular medical practitioner is given access to the records by the convalescents, then only, they will be able to see the patient's records along with the MetaMmask account address of the patient in their dashboard. Figure 2 briefly explains the workflow of the doctors component.

Patient Component

In this part, as depicted in Fig. 3, the patients can register by providing their MetaMask information and sign into their respective dashboards. Later, the patients can upload their medical records, then those files will be uploaded on to the IPFS, and the hash generated will be then stored on to the blockchain. Then, the convalescent can give access to their data to a particular medical practitioner.

Fig. 2 Working of doctor component



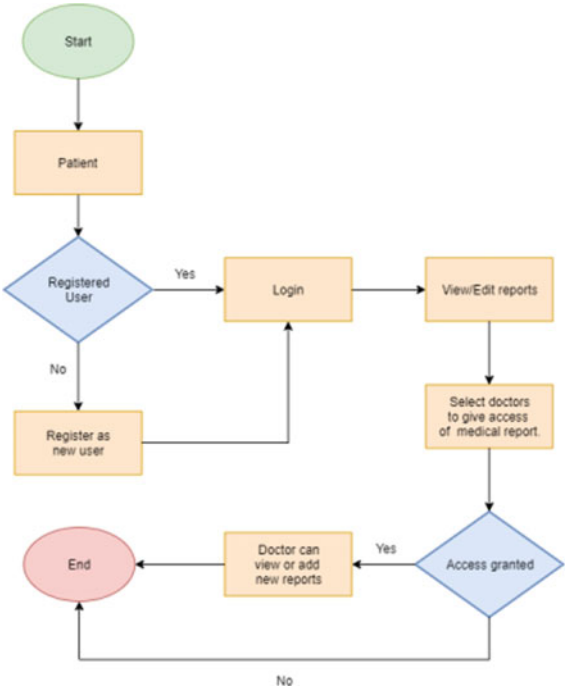


Fig. 3 Working of patient component

Requirements

The basic requirements for the framework to work effectively are processor recommended 1 GHz and RAM of the system as minimum 1 GB, but recommended is to have 4 GB. The storage required is advocated to be 400 MB. It will require the users to have a web browser installed on their systems. Along with the web browser, it is necessary to have a MetaMask account in order to facilitate the transaction cost incurred in the framework. It will also require photographs/media/file access to upload the medical record into the framework.

4 Performance

In this section, we will be evaluating the developed framework. By testing its performance, we can overcome various risks associated with the framework.

To test the performance of framework, we conducted experiments using following arrangements Intel Core i5 CPU @ 1.10 GHz, 8 GB memory, and Windows 10 64 bits.

We developed the framework using solidity a programming language used to write Ethereum smart contracts.

In order to evaluate the performance, we will need some data to gage the performance of the framework; later, we will discuss the parameters or metrics.

Data for evaluation based on whose value various performance metrics will decide the performance:

Transaction Deployment Time (tx1): It can be explained as time when the transaction gets deployed. In Ethereum, smart contracts are deployed using transactions, and this time refers to that.

Transaction Completion Time (tx2): It can be explained as time when the transaction gets completed and confirmed by the chain.

Now, the performance metrics for evaluation is:

Execution Time: It is the difference between confirmation time and execution in the blockchain.

To check if the framework works fine, we would perform all the capabilities of the framework to test the real-life scenarios.

Average Execution Time: The time of execution increases as the number of transactions increases. When a single user uses this framework, the function Add Patient Record takes 37 s, View Patient Record takes 2 min 45 s, and Grant Access takes around 1 min. The time will increase as the number of users in the framework increases.

For every transaction, the fee or cost is associated. In Ethereum, the transaction cost is calculated in ETH which is the currency associated with Ethereum; some smaller units are wei and gwei.

The formula to calculate transaction fees is represented as:

$$\text{Transaction Fee} = \text{Gas Consumed} \times \text{Gas Price}$$

We can calculate the transaction cost by using the recommended figure for Gas Consumed that is 11000 and is 11 gwei for Gas Price.

$$\text{Transaction Fee} = 11,000 \times 11 = 121,000 \text{ Gwei.}$$

And, 1 Ether $= 1000,000,000$ Gwei.

Then,

for 1 Ether, Transaction fee $= 0.000121$ Gwei.

Following Table 1 shows some basic functions that are used in our ReMeDi application. We can distinctly see that each function equates some quota of data and transaction emolument. All the functions that change or modify the smart contact state when they are called usually incur some gas fee. The currency associated with Ethereum are Wei and gwei.

Table 1 Basic functions of ReMeDi application

Transaction	Size	Fee
$Txn_{Register\ patient}$	100 bytes	0.000111 ETH
$Txn_{Register\ doctor}$	100 bytes	0.000103 ETH
$Txn_{Upload\ file}$	324 bytes	0.000252 ETH
$Txn_{Grant\ access}$	164 bytes	0.000398 ETH
$Txn_{Revoke\ access}$	164 bytes	0.000137 ETH

Comparison with different framework

We will be comparing our application with present frameworks in this domain on various parameters without compromising on privacy as well as security.

Scalability

Scalability can be defined as the ability of the system to perform different functionalities, in situations when system storage increases or decreases. After several advancements, blockchain lacks scalability. Our proposed framework works even if the storage volume increases or decreases as we are using IPFS.

As the patient's information is stored in company with the hash of IPFS file, this renders the scalability issue associated with blockchain. Even if patients' data size increases, as the data on blockchain will still be quite less, thus, the transactions will work faster. IPFS uses hash which is stored in a decentralized manner.

Access Control

This framework makes sure that no third party can access the system who is not sanctioned to have access to the system. It provides security using certain safety mechanisms to keep itself secure from third-party incursion.

Content Addressable Storage

Content addressable storage refers to IPFS [33] used for the storage mechanisms. The patient's record is stored on IPFS which generates the hash. It ensures secure storing of patients record.

Integrity

The integrity of any system is defined by how much it can be trusted by its users. It should be genuine and inviolable. Since it is blockchain-based, it does not compromise on this. The information stored is safe and cannot be tampered with by unauthorized access. Patient is the only owner of their data; no second person has the right to modify the data. The use of IPFS locks the security of medical reports.

In Table 2, [35] we have compared our framework ReMeDi with two other frameworks which are FHIRChain and HBasechainDB based on various parameters which include scalability, integrity, access control, and content addressable storage.

Table 2 Comparison of proposed framework with related work

	FHIRChain	HBasechainDB	ReMeDi
Blockchain-based	Y	Y	Y
Scalability	Y	Y	Y
Integrity	Y	Y	Y
Access control	Y	N	Y
Content addressable storage	N	N	Y

5 Conclusion and Future Works

In this paper, we have presented the working of our application-ReMeDi. It provides efficient management and easy accessibility of medical records for hassle-free sharing of health records. It is a boon to all the users who fear the long queues, extended waiting hours, and less secure sharing protocol. Our proposed system is for secure records as well easy access. The system is easy to use and understand. Also, the system ensures to confront the data storage problem as it uses IPFS [32, 34].

The application has transcended the features of existing applications. Later, these features can be added:

- Integrating other actors of the medical field including insurance companies, pharmacists, researchers.
- Integration of consultation to the patient's dashboard.
- Incorporate the payment system.
- Integration of health accessories like Fitbit for health monitoring system.

References

1. Nakamoto S (2008) A Peer-to-Peer electronic cash system. https://www.bitcoin.org/
2. Kostakis V, Giotitsas C (2014) The (A) political economy of bitcoin. tripleC 12(2):430–441
3. Roschin P, Efanov D (2018) The all-pervasiveness of the blockchain technology. Sci Direct 123:116–121
4. Ohno-Machado L, Kim H-E, Kuo T-T (2017) Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 24(6):1211–1220
5. Liu W, Krieger U, Zhu SS (2017) Advanced block-chain architecture for e-health systems. In: IEEE 19th international conference on e-health networking, applications and services (Healthcom), pp 1–67
6. Radanović I, Likić R (2018) Opportunities for use of blockchain technology in medicine. Appl Health Econ Health Policy 16:583–590
7. Angraal S, Krumholz HM, Schulz WL (2017) Blockchain technology applications in health care. Circ Cardiovasc Qual Outcomes 10(9)
8. How blockchain will revolutionise clinical trials. June 2018. [Online]. Available: https://pharmaphorum.com/views-and-analysis/how-blockchain-will-revolutionise-clinical-trials-clinical-trials/. Accessed 20 Nov 2018

9. Alsumidaie M (2018) Blockchain concepts emerge in clinical trials, applied clinical trials. 15 May 2018. [Online]. Available: http://www.appliedclinicaltrialsonline.com/blockchain-concepts-emerge-clinical-trials. Accessed Oct 2018
10. Nugent T, Upton D, Cimpoesu M (2016) Improving data transparency in clinical trials using blockchain smart contracts. F1000Research 1
11. Buterin V (2014) A next-generation smart contract and decentralized application platform. White Paper Ethereum Foundation (Stiftung Ethereum)
12. Mauri R (2017) Blockchain for fraud prevention: industry use cases. IBM, July 2017. [Online]. Available: https://www.ibm.com/blogs/blockchain/2017/07/blockchain-for-fraud-prevention-industry-use-cases/
13. Clauson KA, Breeden EA, Davidson C, Mackey TK (2018) Leveraging blockchain technology to enhance supply chain management in healthcare: an exploration of challenges and opportunities in the health supply chain. Blockchain in Healthcare Today
14. Ajami S, Bagheri-Tadi T (2013) Barriers for adopting electronic health records (EHRs) by physicians. Acta Informatica Medica 2(21):129–134
15. Liu PTS (2016) Medical record system using blockchain, big data and tokenization. In: International conference on information and communications security, pp 254–261, 25 Nov 2016
16. Azaria A, EkblawA, Vieira T, Lippman A (2016) MedRec: using blockchain for medical data access and permission management. In: International conference on open and big data, pp 25–30
17. Engelhardt MA (2017) Hitching healthcare to the chain: an introduction to blockchain technology in the healthcare sector. Technol Innov Manage Rev 7(10)
18. Dagher G, Mohler J, Milojkovic M, Marella P, Marella B (2018) Ancile: privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. ScienceRedirect 39:283–297
19. Chenthara S, Ahmed K, Wang H, Whittaker F, Chen ZS (2020) Healthchain: a novel framework on privacy preservation of electronic health records using blockchain technology. PLoS ONE. https://doi.org/10.1371/journal.pone.0243043
20. Zhang A, Lin X (2018) Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain. J Med Syst 42(140)
21. Fan K, Shangyang W, Yanhui R, Hui L, Yintang Y (2018) MedBlock: efficient and secure medical data sharing via blockchain. Springer
22. Jiang S, Cao J, Wu H, Yang Y, Ma M, He J (2018) BlocHIE: a Blockchain-based platform for healthcare information exchange. In: 2018 IEEE international conference on smart computing (SMARTCOMP), pp 49–56
23. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST (2018) FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 16:267–278
24. Xia Q, Sifah EB, Asamoah KO, Gao J, Du X, Guizani M (2017) MeDShare: trust-less medical data sharing among cloud service providers via blockchain. IEEE Access 5:14757–14767
25. Sahoo MS, Baruah PK (2018) HBasechainDB-a scalable blockchain framework on hadoop ecosystem. In: Supercomputing Frontiers, pp 18–29
26. Omar AA, Rahman MS, Basu A, Kiyomoto S (2017) MediBchain: a blockchain based privacy preserving platform for healthcare data. In: International conference on security, privacy and anonymity in computation, communication and storage
27. Yue X, Wang H, Jin D, Li M, Jiang W (2016) Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst
28. Cichosz SL, Stausholm M, Kronborg T, Vestergaard P, Hejlesen O (2018) How to use blockchain for diabetes health care data and access management: an operational concept. J Diab Sci Technol 248–253
29. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T (2018) Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J Med Syst

30. Liang X, Zhao J, Shetty S, Liu J, Li D (2017) Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In: 2017 IEEE 28th annual international symposium on personal, indoor, and mobile radio communications (PIMRC), pp 1–5
31. Agbo CC, Mahmoud QH, Eklund JM (2019) Blockchain technology in healthcare. MDPI Healthc J 7(2)
32. Gupta S, Sadoghi M (2019) Blockchain transaction processing. Encycl Big Data Technol 366–376
33. InterPlanatery File System (IPFS) [Online]. Available: https://ipfs.io/. Accessed 2021
34. MetaMask [Online]. Available: https://metamask.io/. Accessed 2021
35. Shahnaz A, Qamar DU, Khalid DA (2019) Using blockchain for electronic health records. IEEE Access

Identification of 5G Network Complexity Using SDN

Neetu Faujdar, Ruchi Agrawal, and Ankush Agarwal

Abstract The massive increase in communications technology creates a distributed network, which will be creating a massive quantity of data in billions of network devices. Digital and broadband networks are a significant contributor to this transition which calls for different technological paradigms in a resource-driven world that provides low latency with high efficiency. The full IP-based application launch is planned for 2020. There is, still, no framework for 5G technologies developed, and a standardized 5G specification is being sought out. In the same way, the best choice is changing infrastructure, for instance, software-defined network (SDN). The unified control board is used to separate SDN from the data plane and to monitor the network. This paper provides a state-of-the-art analysis of 5G deployment for the SDN.

Keywords Software-defined network · 5G · Cellular networks · Integrated cellular SDN

1 Introduction

Exponential advancement toward development plus networking creates a hyper-connected ecosystem with millions in linked products. The new cellular innovations like 3G/4G are introducing IP networking that assures to delivery of fast broadband, interactive technology, and a wide variety of facilities with better capacity, versatility in delivery, and lower cost impact. Though, the market for connectivity and the traffic habits of globalization is increasing exponentially, with an unprecedented mobile network increase of 4000-fold over the past decade and membership of 150 million in the first quarter of 2016, with 1.2 billion wired smartphones in the third quarter of 2016 [1].

N. Faujdar (✉) · R. Agrawal · A. Agarwal
Department of Computer Engineering and Applications, GLA University, Mathura, India

R. Agrawal
e-mail: ruchi.agrawal@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_12

Owing to this increasingly growing number of network/submission and interconnection tools, the existing wireless 3G/4G network is not adequately satisfying the planned demands of the 5G wireless system/mobility network for convergence, low latency, and high capacity. The 5G offers user-centered networking, including connections to different devices easier, greater bandwidth, and a 1 ms delay. 5G is an essential tool for the Internet of Things (IoT) framework that connects billions of users as it supports low cost, low power communication machine to machine (M2M), guaranteeing high service quality (QoS). The standardization attempts for 5G are only at an early level, nevertheless, and 5G is expected to be completely realized in "2021 in South Korea, Japan, China," and the USA [2].

To incorporate and operate the 5G wireless networks, current networking and network infrastructure will be reconfigured. That is because of the current 3G/4 G system. We cannot have farsighted stability and cost savings in handling billions of wired devices. In the light of the M2M networking, vehicular access, smart grids, and industrial automation, as anticipated from the 5G IoT network framework, the current spectrum is still inadequate to fulfill planned improvements and efficiency criteria. The software-defined networks (SDNs) and network function virtualization (NFV) allowed innovations are included [3].

The SDN redefines network design by the isolation of the control plane and offers protection for heterogeneous network relationships with constantly evolving complexities through the use of programmable planes. Around the same time, NFV provides many applications effortlessly linked with the network free of the inherent complexities and sharing of capital. SDN network convergence facilitates creativity more efficiently and versatility in delivery resulting in faster connections to emerging IP-based networks [4].

Yet the nuances of 5G, cell network specification and deployment are continuing to be abstracted an active issue in the telecom network. Therefore, the numerous current cellular network structures are envisaged in the sense of SDN and argue for the likelihood of an SDN-based cellular infrastructure reshaping cellular communication. This article highlights different studies that include a cellular SDN-based cellular IoT development approach, in particular, the implementation of 5G in SDN background. Between 2010 and 2016, we review the literature, concentrating on various cellular systems utilizing the freedom of the SDN core router [5].

2 SDN-enabled Wireless World

SDN innovation is a unique methodology into network executives that empower automatically effective organization structure. It is an advanced organization model, in which separation of the control plane from the information plane offers seller freedom, user autonomy, and protection for diverse networks. SDN is normally related to network filtering and functions as a network-level control orchestrate. SDN control airplane is technically a hierarchical organization with a regional network perspective. Network maintenance and setup on the control plane are carried out. The

network supporting the switching system is separated from software and customer facilities. In the data center and cloud infrastructure, SDN is the most commonly used architecture [6].

2.1 SDN Architecture and Protocol

SDN is essentially a three-network architecture. (1) Stage of data plan, (2) stage of pilot and plane power, and (3) stage of operation. Data plane consists of dumb transmitting machines, i.e., routers, which only transmits data on the controller directions. The controller functions as a brain and controls the network from a global network view. Customer specifications are decoupled via applications that, for example, through northbound APIs, are transmitted to the dispatcher. API for RESTful. The manager administers the entire system and has a regional network vision. Many modern sensors, like OpenDaylight, Floodlight, NOX/POX, are on the marketplace from their inception. For incoming data plane flows, the SDN controller determines the law. Northbound interface (NI) API and southbound interface (SI) API enable SDN layers to connect over accessible APIs [7].

The SDN system delivers configuration and adaptive monitoring for the propagation of the meditative state on the regional level with a regional network vision. To satisfy the increasing requirements of customers, SDN will provide high data transmission spectral performance, capital allotment, and management of data for IoT equipment. Many monitoring programs are often meant to operate on the unified device for cell phone networks like the self-organizing networks (SON) program. By tuning the control plane using the SON algorithm, SON generates an efficiency advantage in the RAN [8].

The SDN architecture offers a national network framework for setup and proactive control for the dissemination of the meditative mind. SDN can include fast data transfer, spectral efficiency, capital assignment, and data management for IoT devices to meet the growing requirements of customers [9]. Many control schemes for cellular networks such as the SON software are also configured to run on single devices. SON produces a performance benefit within the RAN system by adjusting the control plane using the SON algorithm [10]. The SDN architecture consists of two types of networks, namely SDN-enabled wireless network and SDN-based cellular network architecture [11].

2.1.1 SDN-Enabled Wireless Network

Wireless networking and SDWN-based optimized SDN wireless networks have also been accomplished with SDN advantages. The wireless network operation is enabled in an economical and fine-grain channeling way via the unified SDN controller design with a strong data rate, with lower latency. The SDWN frequency control is associated with baseband virtualization. A wireless data plane control system is resented to

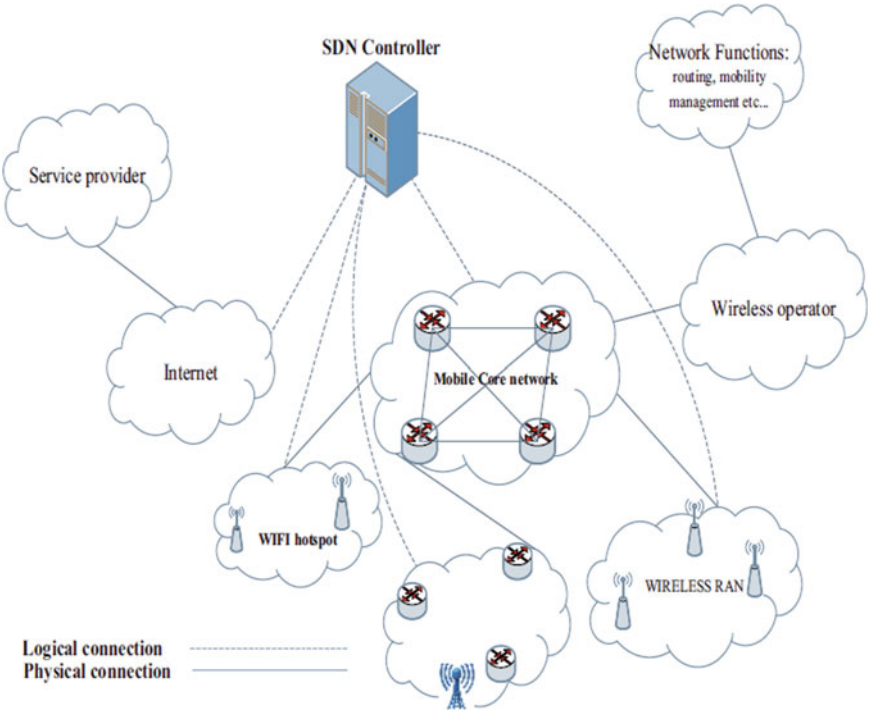


Fig. 1 Mobile wireless network architecture based on SDN [12]

decouple the data plane of the MAC protocol. Data plane management is carried out through proclaimed and flexible programmable device and network protocol (Wi-Fi, LTE) refactoring. To configure the MAC and PHY layers of the data transmission plane, a code application framework is used. Mobile flow is a system with the user modeling the transport grade delivery for the cellular network and the construction of test beds plus the central controller offers security, permission, and the delivery of QoS services. Figure 1 shows a general framework for SDN-based handheld wireless [12].

2.1.2 SDN-*B*ased Cellular *N*etwork *A*rchitecture

QoS for hyper-wired network wirelessly linked over a wide spectrum of networking environments, and cellular network evaluations from first generation (1G) to the fifth level (5G) seek to test the cellular network. A broadband network with a differing spectrum and building a heterogeneous network with various areas including Wi-Fi, LTE, WiMAX, and 3G/4G are the key ingredients of the mobile network. The other approach is to utilize the 4G and 5G networks rather than operator-driven, namely the specifications for infrastructure and standards are different and need production

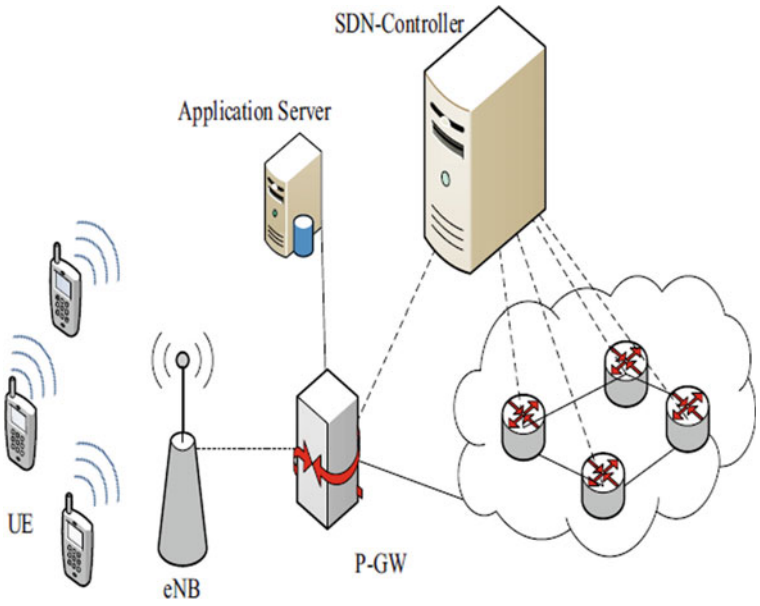


Fig. 2 SDN-based cellular network [12]

and scalable design and installation of the new networks. The hyper-connected 5G network M2M modules involve the convergence of current and new technologies [12].

5G systems comprise the device interface and a varied array of radio access network (RAN) and independent networks. The RAN applications are connected to the outside Internet environment over an Ethernet interface which is handled with a control device policy center. A general cellular network design is presented in Fig. 2.

2.2 Architectural Characteristics of 5G

The expected dreams of 5G organization innovation are centered around a few forecasts, including developed traffic limit, more prominent information rate for every client, diminished inertness, and higher versatile application. Heterogeneous organization sending in 5G is a fundamental plan with a uniting network also.

2.2.1 5G Data Volume and Route

The estimated volume of data from cellular data users is 30.6 exabytes and is averaging 3Mbps link levels up to 2017. This data load is 1000-fold greater than the current level of traffic and 100-fold larger. Therefore, the 5G network's capacity

management aims to revamp the PHY technology and bandwidth performance. 5G manages all oncoming traffic with a single unified system. The SDN theory relates to flow change and control. The software is built with network computation and a genetic algorithm. Flow uses the identical node assets, and this division of labor in this design optimizes the system.

2.2.2 Reduced Latency in 5G

The decreased latency called 1 ms latency, particularly in M2M networking, is an essential feature of 5G technology. The 5G apps will cash local material, e.g., products via EPC, and eNBs decrease reaction time throughout the 5G network. However, eNBs small storage space are not enough to cache and constantly update information, thereby growing the cache impact ratio as opposed to a missed ratio and increasing overall link latency.

2.2.3 Improved Energy Efficiency

As energy-constricted machines, M2M connectivity is generalized. Therefore, energy usage management is important for the higher and longer existence of the energy restraining network. Sleep/wake processes and timing for idle channels may be used as energy-saving nodes.

2.2.4 Key Points of 5G Architecture

In the direction of outlining the essential advantages of the 5G mobile network, dynamically integrated applications that can offer high rate, low latency as well as versatile and configurable software will meet the following criteria. The SDN and NFV definitions will be used for the design of the 5G technology/cellular network.

3 Cellular Network Integration in SDN

Boundary functionality in the telecommunications network is challenging to obtain the intended effect and will actively interfere with radio technology, despite the enhanced sophistication and combined technology. Differentiated control efficiency will improve network output by offering program control and radio control.

3.1 SDN-Based Cellular Network Architecture

Li made the primary suggestion toward an SDN-based mobile network [13]. Wrong and. In s. S, they named it cell SDN, in which the rules for specific LTE network consumers dependent on characteristics are established and the network is managed. Local agents conduct deep packet inspection on-turn in cell SDN and reduce unnecessary controller load [14].

As the transmitter set in which cross-functionalities is challenging to achieve owing to centralized communication algorithm, radio access network (RAN) is managed independently as the network. Different control functionalities can improve network efficiency by offering technology-based radio modulation.

S. is indicated by Soft-RAN. Throughout, Tomovic et al. utilized the 4G LTE network SDN concept. The entire RAN is resumed into the geographical region through a hierarchical control plane. In a grid of three dimensions, a wide base station with centralized power, i.e., size, time and bandwidth slots, is allocated resources. In the field of size, time, and energy, the planner determines to assign capital. Radio aspects/BS wants to tackle disruptions at the local stage [14].

3.2 SDN-Based Cellular Network Management

It is incredibly challenging to set up, redefine, distribute the capital, and also to use the inter-communication model. When controlling the heterogeneous network, SDN has a significant function to play.

D Wu et al. introduced UbiFlow that offers powerful flow management and flexibility administration in urban multi-network via centralized SDN controls. Throughout UbiFlow design, a dynamically dispersed SDN device is used to separate IoT systems into tiny network fragments/clusters. For various data demands, the IoT devices of each partition may be linked to a separate entry point [15].

M. Boussard et al. suggested an IoT system monitoring and control structure based on SDN in an intelligent world. In order to require consumer facilities, the control system, dubbed software-defined LANs (SD-LAN), organizes systems and classes. In order to explore and build a computer topology for device specifications, the Universal Plug and Play (UPnP) and the Standard Network Discovery Protocol (SSDP) architectures are used [16].

4 Proposed Methodology

5G organization comprises of three gatherings of administrations whose necessities ended up being the justification improvement, i.e., massive machine type communication (mMTC), helping bunches of IoT gadgets with inconsistent activity and

correspondences of little information bundles. Expanded circumstances contain a large number of gadgets, robotization of cycles, similar to transportation, shrewd removal of waste, resource following, and so forth. These facts when considered provide a very likely idea that will help us to design the beyond 5G (B5G or so-called 6G) architecture that is seen as ultra-dense networking (UDN) implying a density of hundreds of nodes per square kilometer. To offer improved coverage for every user, a UDN is used which works based on the implementation of moving or static, flying radio access (RA) nodes or ground. To be more precise, the transmission distance is lowered in the UDNs, which results in a better quality in linking and leads to the reusing of frequency to be maximized and increased capacity. For instance, in a crowded event (Fig. 3a), like a trade-fair or a music festival, surplus connectivity is necessary, specifically at remote locations that are not near central areas; typically, the access network nodes are not able to cope with the excess amount of connections by the event watchers.

A strategy to build a network office for helping such occasions can be dispersal of robots, put deliberately for better-quality organization and additional transfer speed

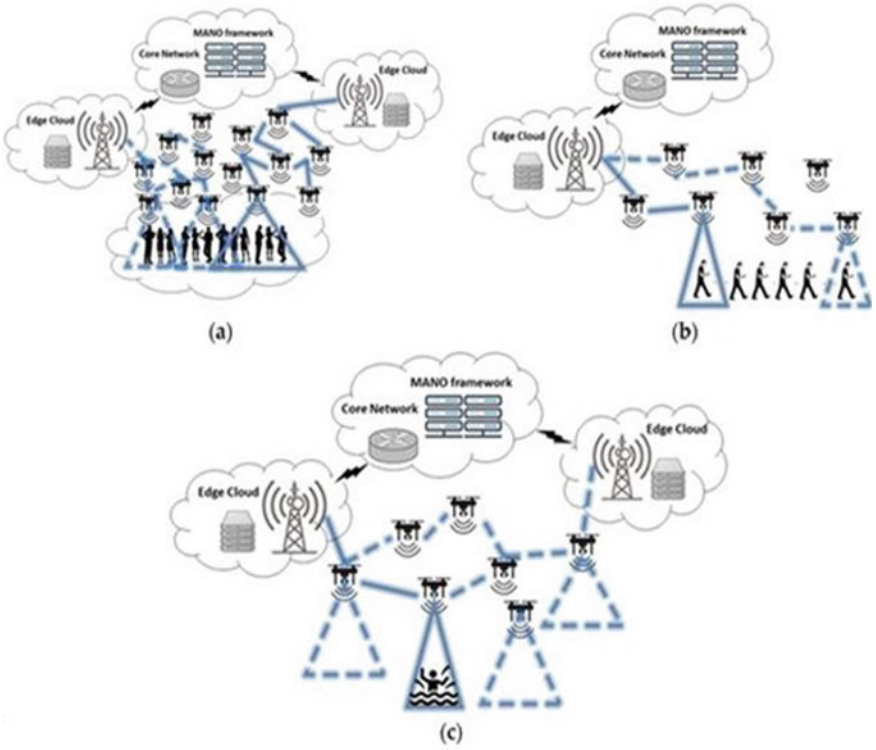


Fig. 3 Fifth-generation network demanding typical use case scenarios: **a** Crowded event; **b** Moving users; **c** Network recovery after a natural disaster

to the space, and simultaneously the transmission of fast associations, the versatile information traffic between center organization and the clients. Moreover, edge content reserving might have the option to facilitate the tension on backhaul connects and further develop client experience.

Looking at a similar situation, several outdoor dwellings of attention for clusters of visitors remain located to some extent outside the range of outdated telco arrangement, comprising of trips on the sea toward coasts, site visits to historical places by individuals, jungle safari trips to view the animals and vegetation, etc. Network coverage may be provided to individual sight-seeing (Fig. 3b), using radio-controlled drones, hovering over places of tourism. The connection among the main network and aerial random access (RA) nodes will need a type of medium to long-range network transmitting equipment. A few RA nodes would be following visitors for the duration of their visit. Lastly, after huge calamities, for instance, hurricanes and earthquakes, network connectivity may be lost, resulting in the isolation of people and leaving disaster management teams without proper communication with the outside world. In these situations, some type of network providing equipment must be set up immediately (Fig. 3c), concentrated on the affected regions, allowing the disaster management teams and individuals in the region to be able to communicate and enabling them access to the Internet or other resources, like voice. The deployed network needs to be mobile for it to follow the disaster management teams, recognizing signals and noises coming from the rubble by using FLIR cameras or other sensors. Consequently, it develops superficial from the overhead that many external challenges can recognize the development of mobile 5G networks, based on real-life usage. The presence of multiple usages with varied abilities leads toward the harmonized practice of connectivity nodes, to provide services and network connectivity for the other facilities, with clashing goals on quality of service (QoS) and freedom of movement for the sustenance of the different networks. Henceforth, the handoff is further complex, so it needs additional study. Finally, impedance organization can be viewed as more troublesome and not reachable by the current clarifications given in the writing. The air of portable hubs, stable BSs, and gadgets/clients produce a complex meddling model which would be further equivalent to a wireless sensor network).

5 Discussion and Open Issues

SDN model may be used on multiple mobile network levels. But, the WSDN currently is still in the early stages, and several concerns remain unresolved. Low-cost systems are resource controlled; those need a service cycle to allow reserve resources communication, long-range communications through several hop communications, data accumulation, and optimization of cross-layers. The proposed implementation is, therefore, not completely implemented using SDN, and a complete design specification has not been developed till date. Few initiatives are genuinely exemplary such

as software-defined radio access network (Soft-RAN), software-defined networking architecture (SoftAir) open road, and mobile flow.

The SDN operator must be able to accommodate compact network node versatility, self-setup, and flow management. The traffic management of every system absorbs bandwidth and thus degrades the spectral efficiency of the cell network. Due to low latency namely the latency of 1 m in the network of 5G routing and flow management is needed through in-network storage that does not yet know what flow will except for the hierarchical control. SDN unified control aircraft can often profit from operation assaults and mid-attack denials. Data protection is a big concern because of the large amount of data generated inside the IoT cellular network. The enormous scope of hyper-connected networks renders it impossible to build the SDN-based wireless network effectively.

6 Conclusion

In 2021, 5G is a technology priority to include modern networking and communications standards that allow an intelligent ecosystem. It affects the way we think of connectivity in an object and increases the quality of life in our community. In any case, 5G conditions have still not been perceived, even though genuine exploration is being taken on to produce an intelligible 5G foundation that can oblige monstrous amounts albeit solid data transfer capacity given by the cell organization. On account of the huge number of associated gadgets and information produced, a cell network needs completely programmable, versatility, assurance, and information preparation.

To meet consumer desires, scripting and unified management are strongly awaited. In this article, the latest mobile network development system for the SDN control plane has been checked.

The core aspects of the 5G network and attempts to reach fast wireless data speeds are being addressed in this article. Finally, SDN will flexibly alter mobile network types and gain consistency in reshaping the wireless infrastructure to meet the required features for the 5G mobile network.

References

1. Sezer S et al (2013) Are we ready for SDN? Implementation challenges for software-defined networks. IEEE Commun Mag 51(7):36–43
2. Hu F, Hao Q, Bao K (2014) A survey on software-defined network and openflow: from concept to implementation. IEEE Commun Surv Tutor 16(4):2181–2206
3. Kreutz D, Ramos FM, Verissimo PE, Rothenberg CE, Azodolmolky S, Uhlig S (2014) Software-defined networking: a comprehensive survey. Proc IEEE 103(1):14–76
4. Nunes BAA, Mendonca M, Nguyen X-N, Obraczka K, Turletti T (2014) A survey of software-defined networking: past, present, and future of programmable networks. IEEE Commun Surv Tutor 16(3):1617–1634

5. Braun W, Menth M (2014) Software-defined networking using OpenFlow: protocols, applications and architectural design choices. Future Internet 6(2):302–336
6. Medved J, Varga R, Tkacik A, Gray K (2014) Opendaylight: towards a model-driven SDN controller architecture. In: Proceeding of IEEE international symposium on a world of wireless, mobile and multimedia networks 2014, pp 1–6
7. Andreev S, Petrov V, Dohler M, Yanikomeroglu H (2019) Future of ultra-dense networks beyond 5G: harnessing heterogeneous moving cells. IEEE Commun Mag 57(6):86–92
8. Hakiri A, Berthou P (2015) Leveraging SDN for the 5G networks: Trends, prospects and challenges. ArXiv Prepr. ArXiv150602876
9. Hicham M, Abghour N, Ouzzif M (2018) 5G mobile networks based on SDN concepts. Int J Eng Technol UAE 7(4):2231–2235
10. Karakus M, Durresi A (2017) Quality of service (QoS) in software defined networking (SDN): a survey. J Netw Comput Appl 80:200–218
11. Ali-Ahmad H et al. (2013) An SDN-based network architecture for extremely dense wireless networks. In: 2013 IEEE SDN for future networks and services (SDN4FNS), pp 1–7
12. Tayyaba SK, Shah MA (2017) 5G cellular network integration with SDN: Challenges, issues and beyond. In: 2017 International conference on communication, computing and digital systems (C-CODE), pp 48–53
13. Jin X, Li LE, Vanbever L, Rexford J (2013) Softcell: scalable and flexible cellular core network architecture. In: Proceedings of the ninth ACM conference on Emerging networking experiments and technologies, pp 163–174
14. Shenker S (2013) Stanford seminar-software-defined networking at the crossroads
15. Wu D, Arkhipov DI, Asmare E, Qin Z, McCann JA (2015) UbiFlow: mobility management in urban-scale software defined IoT. In: 2015 IEEE conference on computer communications (INFOCOM) pp 208–216
16. M. Boussard et al. (2015) Software-defined LANs for interconnected smart environment. In: 2015 27th international teletraffic congress, pp 219–227

Estimation of Dynamic Fault Currents of Microgrids Using Generalized Regression Neural Network

A. Ananth, P. Mohan Babu, and C. Vaithilingam

Abstract Microgrids are paving way to accommodate the distributed generation units (DGs) in the power systems. Besides the technical advantages such as enhanced voltage regulation and reduced power losses, implementation of micro girds faces few challenges also. Differentiating between the normal operating condition and the fault condition is one of the technical challenges. The fault current of microgrids are the function of mode of operation of microgrid like grid-connected or islanded mode. The weather conditions also play a major role in the power generation level of PV or wind energy generation in turn fault currents. As the fault currents vary with respect to many operating conditions, the conventional electromechanical relays or static relays cannot be used for the protection of power system with microgrids. Numerical relays find wide applications in the protection of microgrids. The processor-based relays will be programmed using machine learning or artificial intelligence-based algorithms to estimate the fault current values at a given operating condition. The objective of this paper is to develop a method to estimate fault currents of sample microgrid operates at different operating modes and conditions, with the help of artificial intelligence (AI) techniques. Supervised learning in general regression neural network in which the neural network has been trained with the help of various irradiance values and its corresponding fault current value. The fault currents obtained are compared with the values obtained using conventional Simulink model.

Keywords Microgrid · Neural network · Reinforce learning · Irradiation · Fault current

1 Introduction

Microgrid is an electricity distributed system that can operate in grid-connected mode by connecting to the traditional grid or in islanded mode by disconnection from the central grid. It consists of load and distributed energy sources such as storage

A. Ananth · P. Mohan Babu · C. Vaithilingam (✉)
Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India
e-mail: vaithilingam.c@vit.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_13

devices, distributed generators and some controllable loads. Usage of renewable energy sources such as biomass, bio gas, solar energy and wind energy in microgrids leads to lower environmental impact and increase in stability of supply. When a fault occurs in any part, we can disconnect the faulty part by changing to islanded mode. Due to variable generation and load, the fault current value changes dynamically which necessitate adaptive protection scheme for the microgrid.

Microgrid consisting of distributed generators in the power system emerged in recent years [1]. The conventional unidirectional distributed network can be easily protected as the fault current is not dynamic. The multi-source present in the networks leads to multidirectional power flow in the grid which requires the need of adaptive protection systems as an alternative for traditional protection systems. Presence of distributed generators in microgrids impacts the change in fault current magnitude and direction.

In this paper [2], the coordinated protection using overcurrent relays is considered to be simple. By using real-time values, the artificial neural network (ANN) model is trained to analyse the status of fault conditions in the network. Another ANN model is made with same measurements further to estimate the fault location. To enhance the reliability of operation towards various operational situations, modification of settings of relays is performed by reconfiguration. Accurate and intelligent modification is done to the protection setting using the assistance of estimation models.

This paper [3] proposes a reliable, accurate and adaptive method for protection of microgrids with the help of fault current measurements of the system at different operation modes, different types of loads and with the dynamic nature of distributed generators. In this system, monitoring is done continuously by the central protection unit using communication devices and the fault current settings will be changed based on the instruction of the CPU whenever a fault is detected. Accurate operation is done by estimating the current values under various operating conditions of microgrid instantaneously, and the circuit breaker will be operated accordingly. This makes the protection scheme adaptive, and it can avoid human interferences.

This paper [4] proposes an intelligent secondary controller based on neural network for microgrids to handle faults and uncertainties in the system dynamics. This adaptive neural network controller is based on multilayer perceptron network (MLP).The performance of this controller is analysed based on the simulation results of microgrid with fault condition and disturbance which includes transition from grid-connected mode to islanded mode.

In the journal [5], the author has used an ANN for data analysis. It was found that for the DGA method, ANNEPS takes advantage of both artificial intelligence and human expertise by integrating the results from an artificial neural network and expert system. Further, better results are obtained by using evolutionary programs for optimal coordination of the DOCR system compared to the conventional method for pick up current, TDS and operating time. [6] does a vast discussion on the existing AI techniques for analysis of different network configuration and finds that the existing ANN, fuzzy logic, genetic algorithm, SVM and wavelet-based techniques are successful but are too few. The paper also suggests that there is a huge scope of

research in AI-based techniques which can simplify the complex non-linear systems with proper modifications in machine learning methodology. In paper [7], a method is proposed using multi-agent for the protection of transmission lines.

In [8], the author suggests that field programmable gate arrays can be used for the faster detection and correction of microgrid protection. It is also suggested that modern FPGA is very fast and hence can reduce the time taken for fault detection. Further research is required in the field of microgrids for the development of a sophisticated robust network.

In paper [9], the effects of DER and topological changes on microgrid protection are discussed. A digital relay-based protection scheme is proposed. The proposed method is designed using pre-calculated settings which are automatically updated by the central controller. Relay settings are pre-calculated using off-line studies.

From the literature review, it is inferred that the ability of protective relays to differentiate between normal operating condition and fault current is very much important for adaptive protection schemes. The efficiency of microgrid protection system primarily depends on the ability to discriminate the currents. The accurate estimation fault currents under various operating conditions is essential to develop the AI model for adaptive protection. These factors necessitate the need for a method which comprehensively covers all possible operating conditions of the system and AI-based method which estimates the fault current accurately with lesser computation. This paper proposes an efficient method for the estimation of fault currents using an artificial intelligence-based generalized regression neural network (GRNN) method.

2 Estimation of Dynamic Fault Currents

2.1 Block diagram

As shown in Fig. 1, a dataset consisting of irradiance value and presence/absence of alternator is sent to the trained artificial neural network (GRNN is the model we use, not ANN) model which processes the data and gives the corresponding current value as an output.

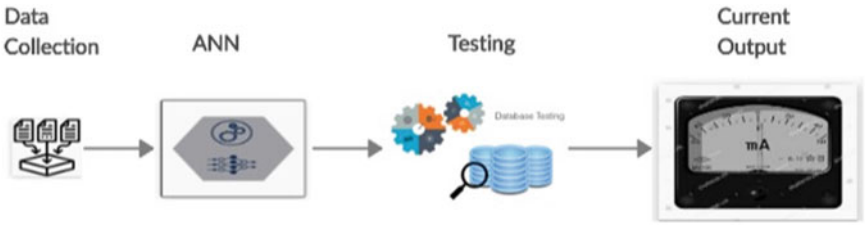


Fig. 1 Workflow diagram

2.2 Model Explanation

For the analysis, the Simulink models in Figs. 2 and 3 were used. The designed microgrid consists of five renewable resources. Islanded mode is given in Fig. 2. The grid-connected mode is given in Fig. 3. The alternator indicates whether the microgrid is connected to the central grid or isolated. If the alternator is present, the system is in the grid-connected mode (Fig. 3); otherwise, the system is in the isolated mode (Fig. 2). The fault current estimation was done first for the microgrid in islanded mode and then for the microgrid in grid-connected mode.

Figure 2 represents model without alternator, and Fig. 3 represents model with alternator analysis that was done for the model which consists of five renewable energy sources. A model was designed for a line voltage of 11 kV. Hence, the inverter's phase-to-phase voltage was set as 16 kV. A sampling time of 2.5 microseconds was chosen for modelling. The model was designed for a load of 10.5 MW. The designing of RLC circuit for the inverter was done using the equations,

$$\text{Cfmax} = (0.05 * P)/(2 * \text{pi} * f * U^2) \tag{1}$$

$$Lf = (0.1 * U^2)/(2 * \text{pi} * f * P) \tag{2}$$

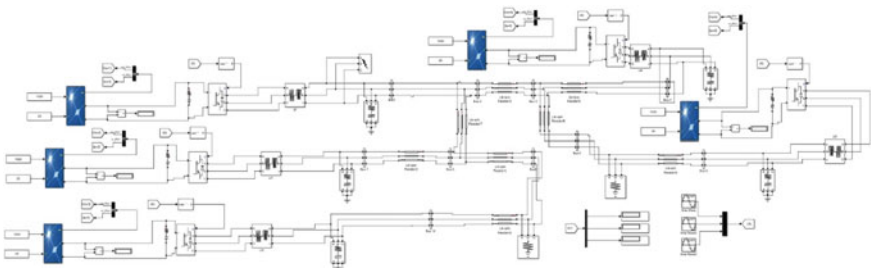


Fig. 2 System without alternator (islanded mode)

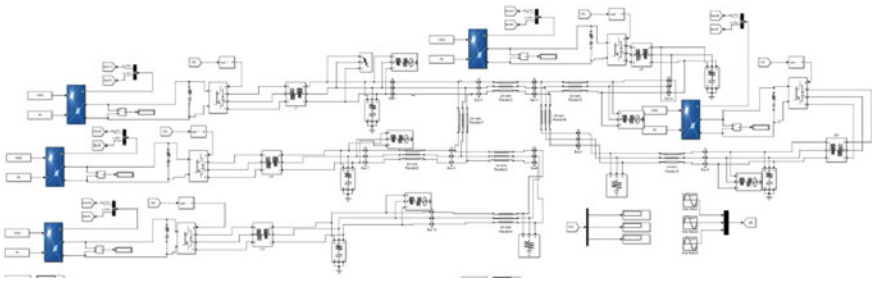


Fig. 3 System with alternator (grid-connected)

$$\text{RLf} = \text{Lf} * 100 \tag{3}$$

Rlf: series resistance.
Lf: series inductance.
Cfmax: parallel capacitance.
P: power rating of the load.
f: 50 Hz.
U: inverter phase-to-phase voltage.

RLC components were used in order to remove higher harmonics and to soothe the waveform from the inverter. A PWM generator was used for switching the inverter.

The parameters for the distribution line were designed considering the characteristics of aluminium cable steel reinforced conductors. The resistance per length, inductance per length and capacitance per length for designing the distribution line were obtained from Characteristics of Aluminium Cable Steel Reinforced Conductors (ACSR) data. Each distribution line was designed for a length of 14 kms. Two purely resistive loads, one of 5 MWs and another of 5.5 MWs were introduced at two locations. Thus, a total of ten busses were introduced. Alternators were attached at buses where bus voltage was dropping.

3 Methodology

Data Collection

Three phase fault blocks were introduced at five locations in the model. The five locations are indicated in Figure 4 as locations: A, B, C, D and E. Fault currents for LG, LL and LLL faults were measured at these five locations. The estimation of fault currents was done for two cases. The first case was with an alternator, the system is in grid-connected state. Distributed generation such as power generation using solar panels, wind energy can be used to generate electricity. In the model, five PV

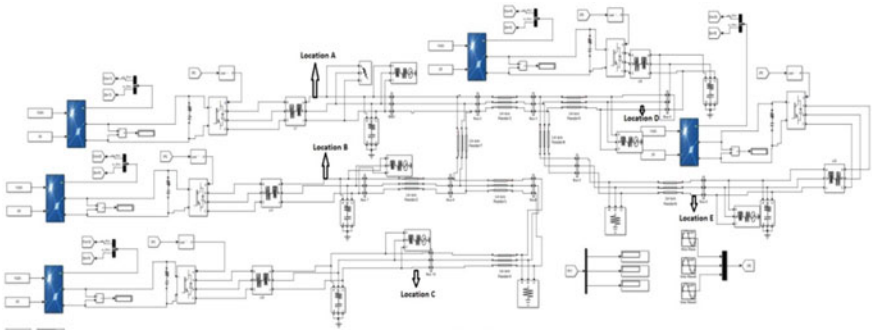


Fig. 4 Model with fault locations marked

sources are connected as distributed generators. All the distributed generators were connected to the grid during the analysis. During the grid-connected state, alternators were connected at all the five locations. The normal current at location A was obtained using Simulink model. Then, a line-to-ground (LG) fault was introduced at location A of the model. The LG fault current was measured. The same procedure was repeated with alternators attached at all the five locations. The normal currents and LG fault currents were noted for all these operating conditions. The power generation of all the PV sources were varied. The different power generation was obtained by varying the irradiance values. The normal current and the LG fault currents at different PV generation were also obtained using simulation models. The irradiance values are varied from 100 to 1200 with a step increase of 100. All the four data was recorded for the set of irradiances. The second case was without alternators connected (islanded mode). The same procedure was repeated for microgrid in islanded mode also. The irradiance values and corresponding values of normal current and LG fault currents are used as training data to develop AI-based protective relaying model.

Training the data and validation

The method of supervised learning was used in order to train the data for neural networks. The value zero indicated the absence of alternator and one indicated the presence. This value and irradiance were used as inputs to train the ANN model. The output is normal current and fault current. The network was trained in MATLAB Simulink. After training the network, it was tested for different values of irradiance values as inputs which were not in the training. The current values thus obtained were compared with the values obtained for the same sets of irradiance using the Simulink model.

Table 1 is the data obtained from the Simulink model, the neural network and the error for normal current obtained for islanded condition.

Table 2 shows the data obtained from the Simulink model, the neural network and the error for fault current obtained for islanded mode.

Table 3 is the data obtained from the Simulink model, the neural network and the error for normal current with the alternator.

Table 4 is the data obtained from the Simulink model, the neural network and the error for fault current with the alternator.

Table 1 Without alternator (islanded mode)—normal condition

Irradiation(W/m^2)	Simulink	GRNN	Error(%)
140	50.84	48.34	4.91
225	53.62	53.08	1.007
360	53.53	55.35	0.32
730	57.66	57.54	0.208
980	58.56	58.566	0.01
1020	58.57	58.567	0.005

Table 2 Without alternator (islanded mode)—fault condition

Irradiation(W/m^2)	Simulink	GRNN	Error(%)
140	197.31	198.24	0.46
225	202.47	202.26	0.10
360	207.92	208.28	0.17
730	213.43	213.38	0.16
980	215.34	215.755	0.19
1020	215.55	215.758	0.096

Table 3 With alternator (grid-connected)—normal condition

Irradiation(W/m^2)	Simulink	GRNN	Error(%)
140	49.15	48.51	1.08
225	50.16	50.42	0.515
360	51.47	51.57	0.193
730	52.84	52.76	0.151
980	53.36	53.34	0.037
1020	*53.348*	53.351	0.0056

Table 4 With alternator (grid-connected)—fault condition

Irradiation(W/m^2)	Simulink	GRNN	Error(%)
140	2089.12	2075.35	0.659
225	2219.58	2206.5	0.589
360	2266.61	2268.64	0.089
730	2314.02	2313.13	0.038
980	2325.45	2325.77	0.013
1020	*2326.02*	2325.95	0.003

4 Result

Using the general regression neural network, the current values for irradiances which were not in the training data was found. The acquired qualities were confirmed by reproducing for a similar irradiance in Simulink. It was found that the error between the values obtained from the neural network and actual values was minimal. It was also found that as the irradiance value increases, the error between actual value and the value from the neural network decreases. That is, a neural system indicated higher accuracy for higher estimations of irradiance. It was also found out that the network was better for grid-connected system than for the isolated system.

5 Conclusion

The protection schemes of power system with renewable energy sources or standalone microgrids require adaptive methods. The protective relays should be able to discriminate between the fault currents and current in normal operating condition. Because same value of current may be a fault current in one operating condition and normal current in some other operating condition. Hence, an adaptive protection scheme is suggested in this paper. The proposed method can be developed and implemented for the protection of a microgrid during any system dynamic uncertainties and disturbances in both the modes of operation. The proposed model is rescalable and replicable to suit any power system irrespective of number of buses or type of energy sources. The simulation results show that the proposed model can be implemented for microgrid protection. The proposed methodology can be deployed in any microprocessor to control the protective setting intelligently during disturbances.

References

1. Choudhary NK, Mohanty SR, Singh RK (2014) A review on microgrid protection. In: 2014 International electrical engineering congress (iEECON), Mar 2014, pp 9–21
2. Lin H, Guerrero JM, Jia C, Tan ZH, Vasquez JC, Liu C (2016) Adaptive overcurrent protection for microgrids in extensive distribution systems. In: IECON 2016—42nd annual conference of the IEEE Industrial Electronics Society, Oct 2016, pp 23–26
3. Peiris WLT, Eranga WH, Hemapala KTMU, Prasad WD (2018) An adaptive protection scheme for small scale microgrids based on fault current level. In: 2018 2nd international conference on electrical engineering (EECon), Colombo, Sri Lanka, 28 Sep 2018
4. Jafari M, Sarfi V, Ghasemkhani A, Livani H, Yang L, Xu H, Koosha (2018) Adaptive neural network based intelligent secondary control for microgrids. In: 2018 IEEE texas power and energy conference (TPEC), 8–9 Feb 2018
5. Geethanjali M, Slochanal SM, Bhavani R (2006) A combined WNN approach for discrimination between Magnetizing Inrush current and Internal fault in a power transformer. IEEMA J, Feb 2006 issue, pp 62–65
6. Ram KK, Nirmala S, Ramesh K, Vishwakarma DN (2013) An overview-protection of transmission line using artificial intelligence techniques. Int J Eng Res Technol (IJERT) 2(1), ISSN: 2278-0181
7. Zhu Z (2018) The protection of power system based on artificial intelligence. In: First international conference on real time intelligent systems: conference paper in advances in intelligent systems and computing, vol 613. Sept 2018, pp 413–422
8. Islam MR, Gabbar HA (2012) Analysis of microgrid protection strategies. In: IEEE international conference on smart grid engineering (SGE‟12) UOIT, Oshawa, ON, Canada, 27–29 Aug 2012
9. Oudalova A, Fidigatti A Adaptive network protection in microgrids. Switzerland Ltd., Corporate Research, Segelhof 1, CH-5405 Dättwil, Switzerland

Enhanced Video Articulation (EVA)—A Lip-Reading Tool

M. Sri Geetha, M. Sujay Sudharshan, S. Sunderesh, P. Aruna, and M. Sanjana Sri

Abstract Lip-reading depends much on the context and the knowledge of language as it does on visual clues and is notoriously challenging. Human lip-reading is a tedious task that requires not only the knowledge about the language but also the visual evidences to approximately or appropriately predict the spoken words. Even the speech recognition in the noisy environment can be made possible through the lip-reading technique. The accuracy achieved through manual lip reading is 40% which can be improvised using AI deep learning techniques. To improve the accuracy of predicting phrases from the low-resolution videos. Enhanced video articulation (EVA) is an articulatory lip-reading technique determining the labial movements of a person. The face of the person is recognised and is segmented for the labia. The process counts on neural networks, AI algorithms comprising many simple computing components coupled together that learn and process information in a way similar to the human brain. As a result, we infer that the deep learning architectures perform similarly to the traditional methods but will improve the word recognition rates.

Keywords Video enhancement · Labial segmentation · Lip reading · Audio-visual speech recognition · SegNet

1 Introduction

Enhanced video articulation (EVA)- a lip reading tool is an AI based tool that is used to predict/generate the words spoken by the speaker from the input low-clarity video. This model can be used to develop a variety of applications, such as a voice replacement for people with laryngectomies. This would also help the people with

M. Sri Geetha (✉) · M. Sujay Sudharshan · S. Sunderesh · P. Aruna · M. Sanjana Sri
Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore—22, India
e-mail: srigeetha.m@srec.ac.in

P. Aruna
e-mail: arunapalaniappan@srec.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_14

hearing disorder to easily recognise the spoken words instead of manual methods of lip reading. EVA does this by converting the movement of the labia to text in real time with enhanced accuracy. This model can also be used to provide quality of experience [1] to the users or the audience of live stream watchers and the online video streaming applications, and the ambiguity and difficulty of the task might be used to replace/overdub genuine speech. Enhancement of the input video does not alter the original size of the file to a greater extent which helps us to overcome the problem of space complexity.

2 Related works on lip-reading

2.1 Video Resolution Enhancement

The study suggested super resolution (SR) methodology that retained more features, resulted in a better super resolved video sequence. Studies on a number of well-known video sequences revealed that the suggested method outperformed both traditional and state-of-the-art approaches.

3 Lip-*R*eading

There are many research works on lip reading using pre-deep learning techniques. Convolutional neural network (CNNs) have been employed in a number of articles to anticipate phonemes or visemes [2–4] from still pictures, rather than recognising entire words or phrases. A viseme is the visual equivalent of a phoneme, "which is the smallest recognisable unit of sound that makes up a spoken word". Petridis et al. train an LSTM [5] classifier using a discrete cosine transform (DCT) and profoundly deep bottleneck features to perceive the entire word.

Wand et al. [6] identified brief words using an LSTM with HOG input features. The persistent use of superficial characteristics is most likely due to a lack of training data in lip reading. However, because the word boundaries must be known ahead of time, the environment is still remote from the real world, as it is with any word-level classification assignment. To determine labelling, it employs a CNN and LSTM-based network and connectionist temporal classification (CTC).

4 Audio–Video Speech Recognition

"Lip reading and audio–visual speech recognition (AVSR) are strongly correlated issues. Using a huge non-public audio–visual dataset, Mroueh et al. conduct phoneme

categorization using feedforward deep neural networks (DNNs). Hidden Markovian models [7] works well with manually created or pre-trained visual features and encodes input images using database records; few researchers used discrete cosine transforms (DCT) and pre-trained convolutional neural networks (CNN) to characterise phonemes; and every one of the three joins these elements with HMMs to group spoken digits or segregated words. As with lip reading, there has been little work put into developing AVSR systems that generalise to real-world situations.

Earlier work can be arranged into two. The main kind utilises connectionist fleeting classification(CFC), in which the model predicts outline-by-outline marks prior to deciding the best arrangement between the casing-by-outline forecasts and yield succession. The output labels are not conditioned on each other, which is a flaw. Sequence-to-sequence models, on the other hand, read the entire input sequence before attempting to predict the output phrase."

The above-mentioned papers predicted the spoken word/phrases from a low-resolution video which was a major drawback; this resulted in low test data accuracy of the models.

5 Dataset Description

For training and testing of EVA, we have created our own dataset. Unlike MIRACL-VC1 dataset [7, 8], which consists of each ten words and phrases spoken by five men and ten women, our dataset consists of both words and phrases uttered by three women and two men. The speakers were positioned against the camera and were instructed to utter the words and phrases as in the table given below. Each dataset instance is composed of a synchronised sequence of colour photographs. Each words and phrases are further decomposed into frames of each 30 and are arranged in the folders for further processing. The EVA dataset contains a total number of 1250 instances (Fig. 1).

Fig. 1 Description of EVA dataset

ID	Words		ID	Phrases
1	As		1	But I do really know
2	Simple		2	But I didn't like you
3	That		3	But I don't know
4	Know		4	It's that simple
5	No		5	It's simple
6	You		6	As simple as that
7	Really		7	As it is
8	But		8	It's time
9	Well		9	All is well
10	Time		10	All are welcome

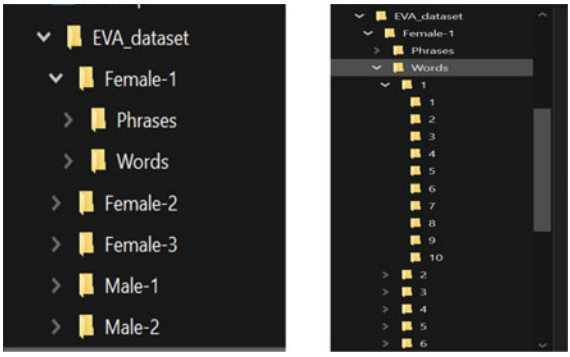


Fig. 2 Structure of EVA dataset and detailed structure of the dataset

The structure of EVA dataset and its detailed structure are given in the Fig. 2. The words and phrases uttered by each of three women and five men are put into their respective hierarchical structure.

6 Proposed model

Enhanced video articulation (EVA) takes low-resolution video with poor or no audio quality as the input. Video super resolution is still an issue which is not solved and is being researched in full capacity. EVA is dissected into three modules, namely video enhancement, labial segmentation, and articulation of the spoken words (Fig. 3).

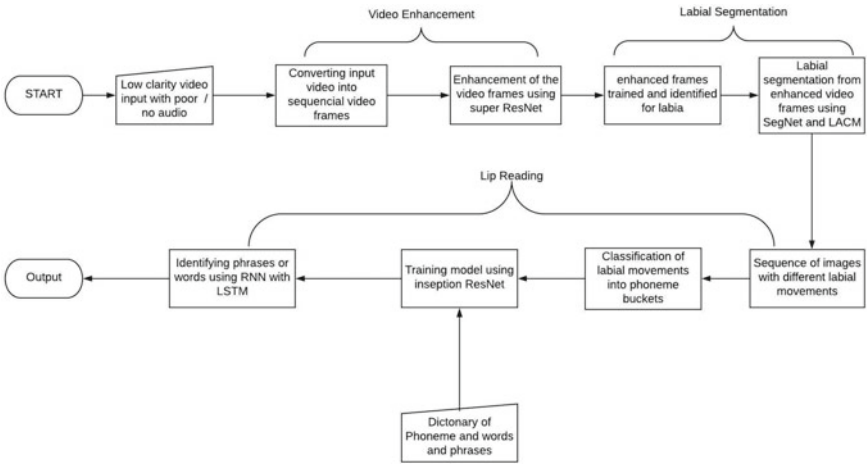


Fig. 3 Architecture of EVA (enhanced video articulation)

7 Video Enhancement

There has been great progress in the field of image resolution in which the image is enhanced from low resolution to high resolution using convolutional neural networks (CNN). The first step is to process the low-resolution videos and convert them into sequential frames. To overcome these problems, we came up with a solution which helps us to convert a low-resolution video to a high-resolution video using deep learning neural networks (DLNN). Hence, the frames of the video are put as the input for our model which uses residual network (ResNet) CNN which enhances the resolution of each frame and stores it into another file. This file is later used as an input file for our final part which converts these HR frames to a video. Thus, the proposed model focuses on enhancing the video at lesser computational cost and consuming lesser time as compared to previous trivial models [8]. The transformed high-resolution frames are then fed into the labial segmentation module where the face is detected using Haar classifier and labial part is cropped using mouth points.

Our model is based on super-resolution convolutional neural networks (SRCNN), but after tonnes of research on various research papers [9], we figured that residual network CNN was a better model as it gives a better image quality with the minimum time complexity. This model uses six residual blocks and 2X upscaling which is a refined version of its parent model SRResNet uses 15 residual blocks and 4X upscaling. After training the model, we tested the model which changed the resolution and the quality to a great extent. Figure 4 depicts the work of super-resolution CNN.

This model can be used in many aspects of work like it can be used to convert low-resolution CCTV footage into high-resolution video which might turn out to be a very important evidence to put criminals behind the bars. It can be used in the field of medicine to convert the videos taken to capture the interiors of the human body with tiny cameras (Figs. 5 and 6).

$$Sum = X(C) + X(A) - X(B) - X(D) \tag{1}$$

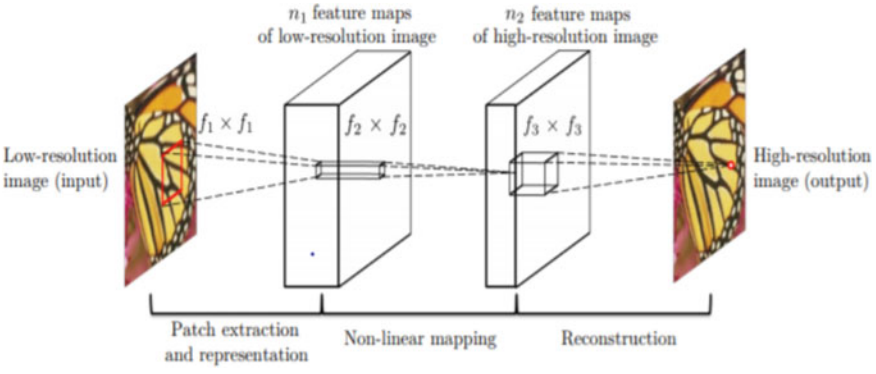


Fig. 4 Super-resolution CNN

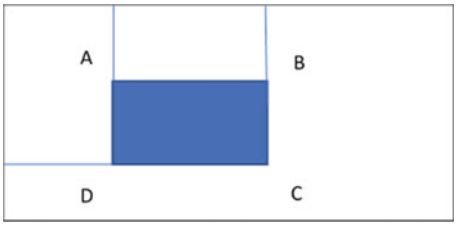


Fig. 5 Finding the sum of the shaded rectangular area

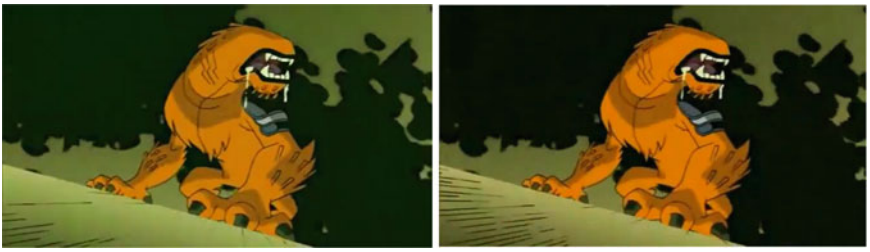


Fig. 6 Input image and enhanced image

where A, B, C, D are the points belonging to the integral image X.

8 Labial Segmentation

"The precision of the segmentation findings has a direct impact on the identification rate in visual lip-reading systems, hence labial segmentation is crucial. Labial segmentation strategies include the MAP-MRP framework, a clustering method, and an active contour model (ACM). Because of its benefits over traditional image segmentation approaches, the LACM model is adopted. ACM can acquires object boundary precision down to the sub-pixel level, energy minimization and accurate outcomes.

Active contour model (ACM) is a vital part image segmentation and computer vision. ACM is classified as edge-based model and region-based model. An image gradient is frequently used in edge-based models to drive the active contour to migrate towards the object's intended bounds. The suggested models were not employed for labial segmentation since they had an incomplete convergence problem due to weak object boundaries and picture noise. The active contour model's scope is limited to a particular region, which limits the effects of unrelated elements. The localised area-based active contour model (LACM) is proven to be capable of creating superior segmentation results where ACM fails.

The labial segmentation is done with the help of a local active contour model, which generates the initial contour automatically. The evolving curve in LACM divides the local neighbourhoods into two areas: the local region that lays inside and

the local region that lays outside. The limited energy for advancing and separating would then be able to be done. However, with LACM, incorrect parameters such as high radius or long-evolving curves cause incorrect prediction. According to studies, the labial is generally of an elliptical region. As a result, the labia can be approximated by various elliptical outlines based on their particular structure. The min-bound oval shape as the underlying developing bend is of critical significance to separate the labial shapes. To track down the base bouncing oval of labial district, the identification of labial corner specks is required. In particular, $X(x, y)$ address a pixel esteem at arrange (x, y), m and n are the most extreme upsides of lines and sections. The left corner, right corner, upper corner, lower corner is signified as L a, L b, V a, and V, individually (Fig. 7).

The steps involved in the extraction or segregation of labial contour are,

- Locate the labial area;
- Acquire the min-bounded oval shape;
- Proceed and develop iteratively;
- Segregate the labial contour area.

SegNet core trainable segmentation engine comprises of networks that actually encodes and decodes and a layer that does the classification pixel-wise. The decoder network maps the low-resolution encoder feature to full input resolution feature. SegNet's distinct quality is that it provides us with the low-resolution feature map that serves as input and it has been decoded too. To conduct non-linear up-sampling, the decoder leverages pooling indices obtained in the matching encoder's max-pooling step."

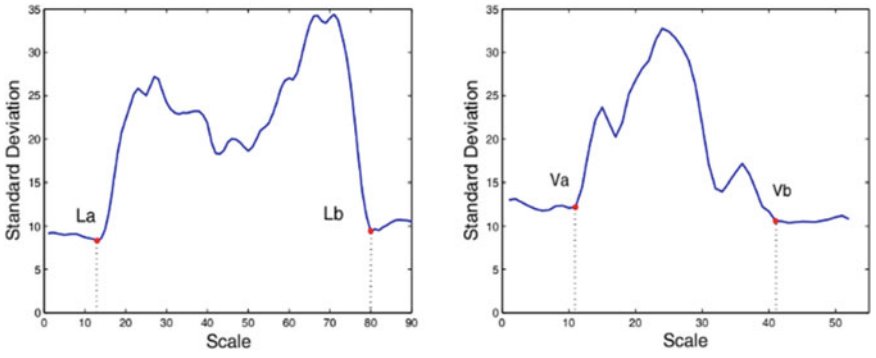


Fig. 7 Standard deviation of column and rows

9 Articulation

The feature-extracted enhanced frames are applied to RNN with LSTM architecture to predict the words in the respective frames. RNN and LSTM, as well as their variations, require time-based sequential processing. Unlike neural turning machines, its design saves all prior representations in memory. This can be inefficient: consider saving the representation of each frame in a movie; most of the time, the representation vector does not vary from frame to frame, thus we are holding an abundance of absolutely similar pictures.

RNN and LSTM necessitate memory-bandwidth-bound computation that will restrict the relevance of neural networks solutions. Long short-term memory (LSTM) with long storage unit acts as an agent that used reinforcement to educate its parameters and to memorise the letters for predicting the term.

The articulation part is structured into three:

- Labial localisation system to locate the labia in the digital input.
- Feature extraction system, which evaluates the appropriate labial features.
- Classification system, which maps feature vectors to terms.

Labial localization: We took an unconventional approach to labial localisation. The primary components were computed using a series of digitised labial pictures. The fixed portions of the image were compressed using these primary components. The labia will be in the region with the least amount of information loss since the major components were fine-tuned for labia.

Feature extraction: A normalised rectangle comprising the labial segment is fed into the feature extractor. For the rectangular input frame, principal component analysis is employed to build a feature vector. A single-layer feedforward neural network is used to estimate the main components.

Translation to words: Words (visemes) must be categorised from the feature extractor's feature vectors. A time-delay neural network was utilised to categorise the feature vectors in order to incorporate temporal features into the classification process.

10 Result

Recurrent neural network with long short-term memory (RNN with LSTM) is used, and the model is fine-tuned with $L2$ regularisers with value 0.1 and 0.5. Optimizer-SGD with learning rate 0.005 is set. The model shows result in 80th epoch with training accuracy as 93% in the end for $L2 = 0.1$. The validation accuracy gets better result (51%) when epoch is 21 and $L2 = 0.1$.

In Fig. 8, Female-I uttered the phrase "But I do really know". Our model predicted the phrase as "But I don't really know" (Fig. 9).

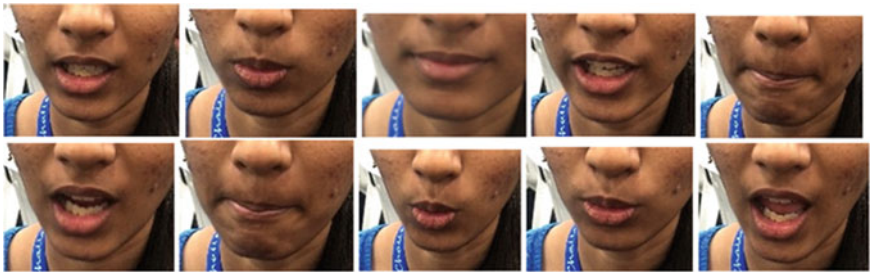


Fig. 8 Clip that contains the phrase "But I don't know" by Female-1

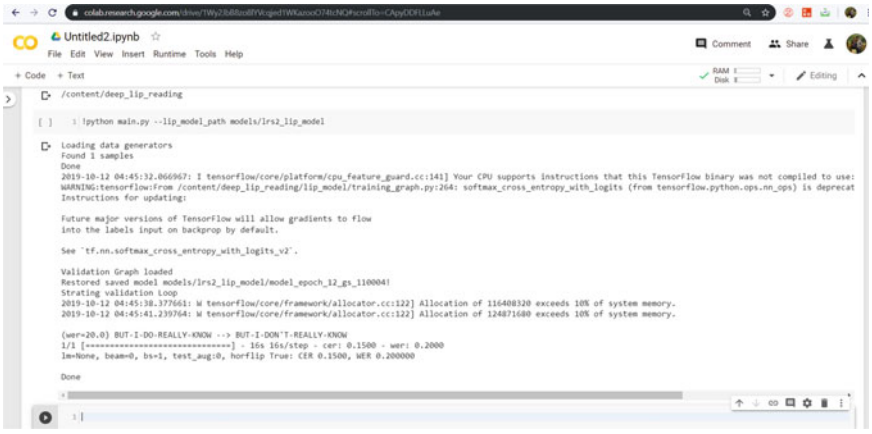


Fig. 9 Predicted output

11 Application

This model can be used to create different application that can be proposed such as an alternative voice for laryngectomised patients. This would also help the people with hearing disorder to easily recognise the spoken words instead of manual methods of lip reading. The output of EVA can be taken for spying purposes, that is, to generate the words spoken in an audio-less video from CCTV cameras with low/poor quality. This can be integrated with the live videos to generate subtitles for that instance accurately. This acts as a valuable communications tool for deaf and hard of hearing people.

12 Conclusion

In this work, we proposed a novel feature approach utilising an enhanced version of SRResNet CNN to convert low-resolution to high-resolution videos. Our experimental results demonstrate a supervised learning approach that recognises ten self-curated phrases. EVA (enhanced video articulation) is a model which performs the labial articulation of enhanced video. SegNet and LACM are used to extract the labia from enhanced frames. These enhanced frames are applied to RNN with LSTM architecture to predict the words in the respective frames.

This work leads to a fundamental understanding of existing models and overcoming their drawbacks. Future work can be developed that is independent of speakers' phoneme by creating larger dataset, increasing the number of phrases uttered by the speakers.

References

1. Desai D, Agrawal P, Parikh P. Soni MP (2020) Visual speech recognition. Int J Eng Res Technol (IJERT)
2. NadeemHashmi S, Gupta H, Mittal D, Kumar K, Nanda A, Gupta S (2018) A lip reading model using CNN with batch normalization. In: 2018 Eleventh international conference on contemporary computing (IC3)
3. Barkowsky M, Sedano I, Brunnström K, Leszczuk M, Staelens N (2014) Hybrid video quality prediction: reviewing video quality measurement for widening application scope. Multimedia Tools Appl 74(2):323–343
4. Bear HL, Harvey R (2016) Decoding visemes: improving machine lip-reading. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 2009–2013
5. Chung JS, Zisserman A (2016) Lip reading in the wild. In: Asian conference on computer vision. Springer, pp 87–103
6. International Phonetic Association (1999) Handbook of the International Phonetic Association: a guide to the use of the International Phonetic Alphabet. Cambridge University Press
7. Wand M, Koutn J et al. (2016) Lipreading with long short-term memory. In ICASSP'16, pp 6115–6119
8. Ephrat A, Peleg S (2017) Vid2speech: speech reconstruction from silent video. arXiv preprint arXiv:1701.00495
9. Mei Y, Fan Y, Zhou Y, Huang L, Huang TS, Shi H (2020) Image super-resolution with cross-scale non-local attention and exhaustive self-exemplars mining. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5690–5699

Technology Applications in Education, Applications of Accelerometer Wearable Sensor and Internet User Gender Prediction—and FDI Attraction for Digital Sector

Dinh Tran Ngoc Huy, Nguyen Thi Hang, Trung-Hieu Le, Nguyen Dinh Trung, and Nguyen Thanh Hoang

Abstract According to statistics by the Department of Telecommunications (Ministry of Information and Communications) in November 2016, the total number of broadband Internet subscribers' fixed broadband reached more than 9 million subscribers and mobile broadband subscribers reached more than 12.6 million subscribers. Besides, according to the statistics of "wearesocial.net," in January 2015, Vietnamese people Nam is ranked fourth in the world in terms of time spent using the Internet with 5.2 h per days, just behind the Philippines at 6 h, followed by Thailand with 5.5 h, and Brazil is 5.4 h/day. So we use quantitative methods and use methods of meta-analysis, comparative statistical methods, survey methods, qualitative and quantitative analysis methods to conduct a study to explore the advances of technology applications into modern education because applications such as accelerometer wearable sensor or Internet user gender prediction can be applied on to online teaching or e-learning process well. In addition to that, FDI attraction is helping provinces in Vietnam such as Thai Nguyen to develop local economic strengths,

D. T. N. Huy (✉)
Banking University HCMC, Ho Chi Minh City Vietnam—International University of Japan, Niigata, Japan

N. T. Hang
Thai Nguyen University, University of Information and Communication Technology, Hanoi, Vietnam
e-mail: nthang@ictu.edu.vn

T.-H. Le
Dai Nam University, Hanoi, Vietnam
e-mail: hieult@dainam.edu.vn

N. D. Trung
National Economics University (NEU), Hanoi, Vietnam
e-mail: trungnd@neu.edu.vn

N. T. Hoang
University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh, Vietnam
e-mail: hoangnguyenfir@hcmussh.edu.vn

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_15

creating more employment and income for the public in the globalization era, expand productivity and quantity for digital sector.

Keywords Technology applications · Digital transformation · Digital sector · Wearable sensor · Internet user gender

1 Overview

First technology application in education: In online teaching or e-learning, especially, during COVID-19 time, we can use some computer programs to help to improve education services for learners. Among these, solutions and applications are prediction of gender of internet users. The sex determination method to classify Internet users is implemented performed using machine learning techniques, using user information of known gender and information about their Web browsing history to train the receiver know the gender of other users when we only know the access history of the site and directory data of interest to the person.

Second technology application direction: In a few recent studies, to observe human activities to increase recognition efficiency, people are simultaneously observed by environmental cameras while still wearing accelerometers as studied by Tran [1]. Even so, such a multimodal information acquisition system still faces difficulties such as installation cost, field of view, and difference between viewing angles. Recently, image sensors have been widely used to track the journey/daily activities of people in a number of sports and health monitoring applications. In these studies, the camera could be worn in several different locations on the body such as on the chest, wrists, and head. The combination of wearing multiple sensors has also been conducted by a number of research groups, but currently, the methods are not geared toward user convenience (each sensor is worn in one position), so it is only suitable for specific applications (care and monitoring of health progress and assessment of recovery) that are not yet aimed at general users.

Next, Nguyen Phu Tu and Huynh Cong Minh tested the two-way relationship between FDI and economic growth in 64 provinces and cities of Vietnam. The results of the study show that the relationship between FDI and economic growth manifests itself in many aspects, including positive contributions to economic growth, job creation, technology transfer, and negative effects related to social problems. On the other hand, the way investments are allocated will yield different results. The impact of FDI is also evident in the mobilization and transfer of technology to modernize and improve production capacity and technological level for the investee country itself.

2 Research Approach and Methods

First of all, this study conducts SVM method, gender determination using data from Twitter messages and regression method in order to address implications of technology applied in education based on two applications: first, technology application in education with prediction method of Internet user gender (learners gender from distance) and second, technology application with accelerometer wearable sensor. Several papers are the base of this work including: study of accelerometers by Tran [1] and a study of accelerometer wearable sensors by Avci et al. [2].

The article conducts an overview of the research works of domestic and foreign scholars, then from there, evaluates the effectiveness of FDI attraction activities to promote economic growth in the digital transformation period.

Authors use meta-analysis aims to synthesize collected and searched documents to systematize theories related to technology applications, combined with survey, synthesis, and inductive methods.

Authors also analyze the fundamental theoretical issues as the basis for making judgments and analyses. And compared to existing similar studies, this work will be different in an aspect of technology application in education with prediction method of Internet user gender (learners gender from distance) to help tailoring e-learning courses, and stated the need for FDI attractions and its efficiency in private sector. There will come the novel of this work.

3 Research Results

3.1 Technology Application

Technology application in education with prediction method of Internet user gender (learners' gender from distance).

Use available data PAKDD'15 is provided by FPT Joint Stock Company (http://www.fpt.com.vn) now construct dataset from existing unnormalized actual data PAKDD'15 for a number of users, using the SVM machine learning technique in Chap. 2, and some tool to give the rate and accuracy of sex prediction method based on access history.

Gender determination using data from twitter messages using regression method.

(a) **Introduction**

Gender determination using data from Twitter messages is a method of classifying each comment by feature based on the comment content using the regression method. In the first step, from the raw data set of Twitter comments collected by topic, we proceed to preprocess Twitter special characters, close duplicates, acronyms, languages, slang, emoticons, and semantic networks.

(b) **Ideas**

Reading someone's Twitter content, in some cases, one can somewhat guess the user's gender. For example, can you know the gender behind the following twitter user?

I LIKE PLAYING FOOTBALL < 3

Regression is a supervised learning method in machine learning. The main goal is to find the relationship between the features of a certain problem. More specifically, from a given data set, we build a model (equation, graph) that best fits the data set, showing the trend of variation and the relationship between the features. When a new data sample comes in, based on the model, we can predict the value of that data sample.

For example, we need to predict the gender of a Twitter based on the content and writing characteristics of that Twitter. Thus, we need to find the relationship between gender depending on content and writing characteristics. Based on the dataset (assuming collecting the content, writing characteristics, and special characters of 100 Twitter users), we build an equation $y = \theta 0 + \theta 1 \times 1 + 2 \times 2$ where y is the dependent gender $\times 1$ (content) and $\times 2$ (written feature). When there is a new sample of data for a new user, simply applying such an equation will predict the gender of that person.

We see that the equation $= \theta 0 + \theta 1 \times 1 + \theta 2 \times 2$ is the equation of the plane in 3D space. Similar models to the equation of a line, the equation of a plane is a linear model. Linear regression is a simple model in a regression problem in which we use lines, planes, or linear equations in general to predict trends in data. Solving the linear regression problem is to find the parameters $\theta 0$, $\theta 1$, … to determine the linear equation (Fig. 1).

One of the problems encountered while running a linear regression model is overfitting. Overfitting is a problem that occurs when the model we create tries too hard to match the samples in the training set. Although the model matches the training

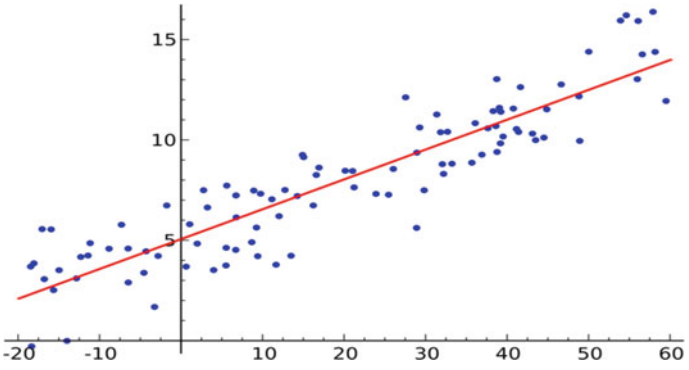


Fig. 1 Example of linear regression *Source* Trung-Hieu Le thesis, Tran [1]

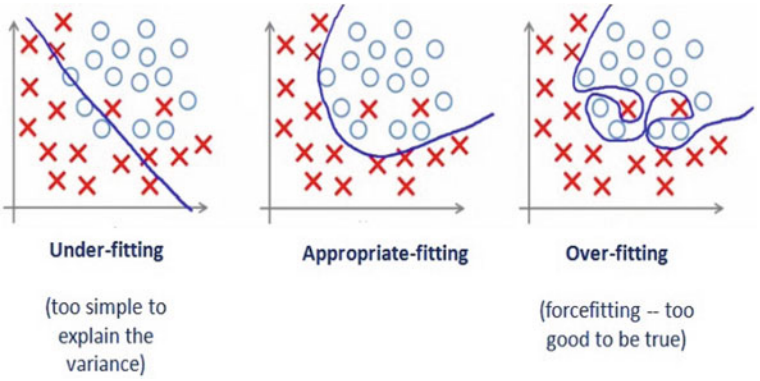


Fig. 2 Matching process *Source* Trung-Hieu Le thesis, Tran [1]

samples, it does not show the trend of the data, leading to the model being true only for the values in the training set and completely wrong for the test values (Fig. 2).

Too much problem matching usually occurs when our Twitter dataset has many features but few data samples. For example, we want to create a model with the form of a straight line, that is, we need two features $\times 1$, $\times 2$ (coordinate features in the plane), but only one data sample is represented as a point. To define a straight line requires at least two points, and if there is only one point, there are infinitely many models that fit the data sample, but only one of them is really true to reality.

Implied practical meanings

Based on prediction of Internet user gender, we can know male or female prefers which courses or educational programs, then we can adjust our course outlines and designed program to meet their demand.

3.2 Technology Application

Technology application with accelerometer wearable sensor.

Moreover, in education and physical exercises, we can think of using accelerometer wearable sensor.

(a) Wearable sensor for human activity recognition problem

Accelerometer wear sensor: Accelerometer is the most popular wearable sensor because of its compact size, easy integration on existing devices (phones, necklaces, bracelets, belts, etc.) at low cost. Many studies using accelerometers have been deployed in various fields such as healthcare (monitoring, diagnostics, rehabilitation assessment, care for the elderly and children), monitoring and surveillance,

family (support for dementia patients), and sports (assessment of energy consumption, performance assessment of sports players). Avci et al. [2] presented an overview of studies using accelerometer wearable sensors.

Due to specific applications, accelerometers can be mounted in one or several locations on the body. Common wear locations for accelerometers are thighs, hips, chest, wrists, arms, ankles, soles of shoes, or at all major joints of the body. The number of body-mounted accelerometers can range from 1 to 48. The activities measured from the accelerometer can take place indoors or outdoors, from cooking activities, personal activities (brushing teeth), walking and climbing stairs to running activities, playing ball/sports, and climbing. For example, in an assistive study of the elderly by Stylianos Paraschiakos shown in Fig. 3, in which accelerometers were mounted at various locations.

The limitation in observing the human hand due to the camera's viewing angle: If the camera is worn on the chest or above the person's head, the angle of view of the hand and the scene in front of the person's face is completely natural, almost the same like the human eye. However, when designing to integrate wearable sensors, the camera viewing angle is limited. It is more difficult to observe the hand wearing device, and the viewing angle of the hand as well as the objects in front of the human eye is not directly. Furthermore, since the human hand is always in motion even when the person is stationary, the signal and the change in angle of view will be very large. This is perhaps the biggest challenge of hand-mounted sensor-based gesture recognition.

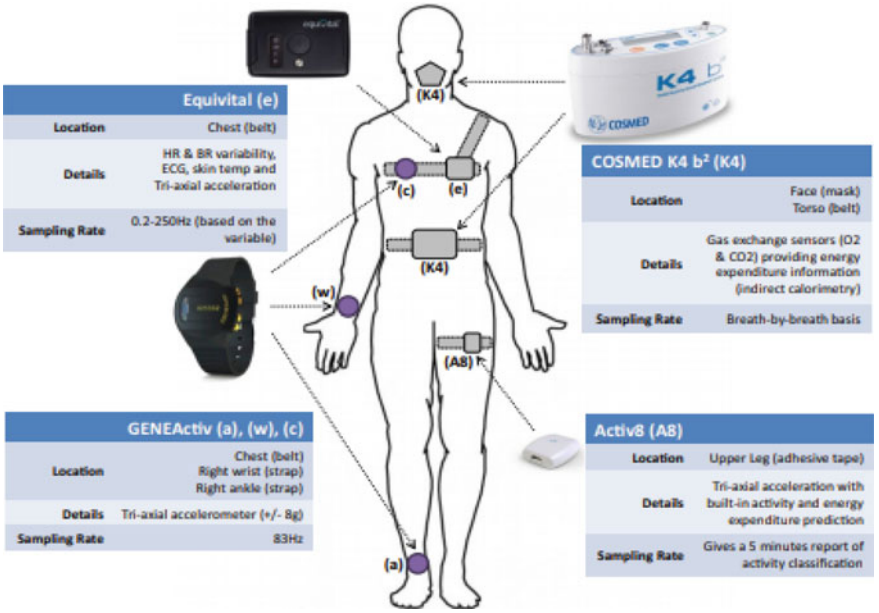


Fig. 3 Some common accelerometer installation locations *Source* Trung-Hieu Le thesis, Tran [1]

Correlation between accelerometer signal and image signal: When the hand performs a certain gesture, the image signal and the accelerometer will be collected simultaneously to serve the recognition process. If the accelerometer measures the movement of the forearm (where the sensor is worn), the visual signal tries to capture the shape and movement of the hand and the front view, in which the movement of the hand and forearm can be completely different. It is not simple to merge a model of gesture representation from multimodal signals in each mode that express different meanings.

3.3 Assessing the Impact of FDI on Industry Growth

Industry growth rate

Industrial production value in 2004 (at current prices) reached 7.7 trillion VND; by 2020, it will reach 879 trillion VND, 114 times higher than the industrial scale in 2004. In which, the processing and manufacturing industry always accounts for a high proportion and is the driving force for growth (Fig. 4).

Compared to the whole country in 2020, the scale of industrial production value of Thai Nguyen province ranks fourth, after Bac Ninh, Binh Duong and Ho Chi Minh City. Compared with other provinces in the region, Thai Nguyen province's industry is the highest and always ranks first. Average annual industrial growth rate: Industrial production value in 2004 (at 2010 constant prices) reached 10,671 billion VND; by

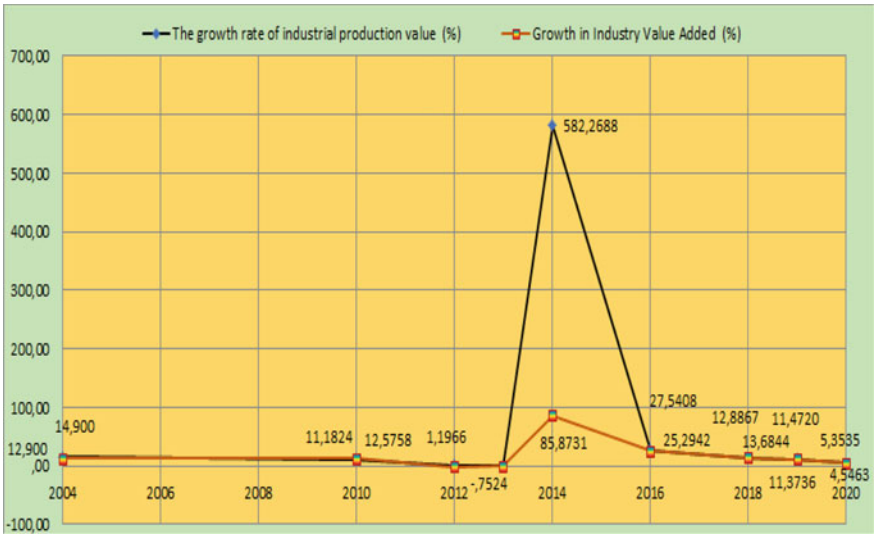


Fig. 4 FDI industry growth rate. *Source* Author's summary from Thai Nguyen Provincial Statistics Office

2020, industrial production value reached 783,619 billion VND, 73.4 times higher compared with 2004. On average, in the period 2004–2020, the industrial production value increased by 29.8%/year; period 2006–2010 increased by 15.45%/year; due to outstanding achievements in attracting FDI investment of the province in the years 2013–2015, the period 2011–2015 achieved the highest growth rate of 72.2%/year; period 2016–2020 reaching 15.8%/year.

4 Conclusions

Huy [3] stated E-learning solutions have changed the way of teaching methods and learning styles of students. And in modern education, we need innovation approaches in teaching methods at our universities and colleges.

This study has presented two approaches: implications of technology applied in education based on two applications: first, technology application in education with prediction method of Internet user gender (learners' gender from distance) and second, technology application with accelerometer wearable sensor. Based on prediction of Internet user gender, we can know male or female prefer which courses or educational programs, then we can adjust our course outlines and designed program to meet their demand.

Next, the 4.0 technology revolution and the digital transformation context create many advantages for attracting FDI into Vietnam. The 4.0 technology revolution will create many challenges for the Vietnamese economy as well as attracting FDI into Vietnam. Because the current structural status of the economy is still inadequate, the transition is slow, and the growth is not sustainable. The efficiency of investment capital (ICOR coefficient) of state-owned enterprises is still far behind that of the private sector. The imbalance and lack of mechanism in policy have led to a large difference in the ICOR coefficient of the FDI sector compared to the non-state sector.

Acknowledgements Thank you editors, friends and brothers to assist this publishing.

References

1. Tran TH et al (2018) A multi-modal multi-view dataset for human fall analysis and preliminary investigation on modality. In: 24th international conference on pattern recognition (ICPR). IEEE
2. Avci A et al (2010) Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: a survey. In: 23th international conference on architecture of computing systems 2010. VDE
3. Huy DTN, Le TH, Thang TD, Hoa NT, Hue LT (2021) Discussion on E-learning solutions for students–and issues of technology application in classroom. Des Eng 7:11432–11443

The ICT Impact on Bank Performance: The Case of Vietnam

Toan Linh Vu Le and Duy Khanh Pham

Abstract Bank digital transformation boosts the development of information and communication technology (ICT) in the banking and financial systems worldwide. This study investigates the impact of ICT development on the performance of the banking system in a frontier market. For the period from 2009 to 2020, the empirical evidence is based on balanced panel data of 39 banks in Vietnam. The result shows that bank profitability and ICT index have a significant positive nexus. This beneficial impact is particularly evidenced in the bank's digital transformation. Hence, the beneficial impact on the return on IT infrastructure investment surpasses the original installation expenses. The findings suggest that ICT advances help the banks perform better when transforming from traditional to digital systems.

Keywords ICT · Banking · Performance · Financial stability

1 Introduction

Banking and finance are the backbones of any country and play a vital role in economic growth [1]. The main reason for the countries' slow social, economic, and even cultural progress is their inability to recognize or develop appropriate technology and use it in productive activities. The effects of ICT (information and communications technology) are to improve productivity and economic growth at both firm and country levels [2–4]. The coverage of the ICT term is vast, including any communication means (such as radio, television, mobile devices, computer, laptop, and satellite systems) and various internet services and applications that go along

T. L. V. Le (✉) · D. K. Pham
Faculty of Banking and Finance, Van Lang University, Ho Chi Minh City, Vietnam
e-mail: toan.lvl@vlu.edu.vn

D. K. Pham
e-mail: duy.pham@ueh.edu.vn

School of Banking, University of Economics Ho Chi Minh City (UEH), Ho Chi Minh City, Vietnam

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_16

with them (such as Web conferencing and online learning) [5]. ICT is suitable for different sectors of the economy. In particular, the banking system has widely applied technology innovation since the 2000s when the Internet grew popular after nearly ten-year development. At that moment, the booming of electronic banking was one of the most significant marks [6]. It is not hard to find an example of ICT products used in the banking sector, must be mentioned as automated teller machines (ATMs), chip cards, phone banking, magnetic ink character recognition (MICR), electronic funds transfer (EFT), electronic data interchange (EDI), and online banking. For the banking business, ICT applications brought up new markets, new goods, new services, and more effective delivery platforms, including online, mobile, and Internet banking [7]. Thus, we can observe considerable efforts from various banks to keep in track with modern technology application trends to enhance their products' features, utilities, and quality, aiming to not only to create a more significant competitive advantage but also to save labor and delivery costs, then grow banking profitability. Digital transformation has provided banks with plenty of opportunities to reach and offer services to many customers, promoting comprehensive financial development [8].

In this study, we looked at the function of the ICT index in the banking system. More specifically, we want to see if ICT is associated with bank performance or improves information distribution. Research results indicate that bank profitability and ICT index have a positive nexus. This beneficial impact is typical in the banking system's technological endowment. As a result, the favorable impact on the return of IT infrastructure investment surpasses the initial installation costs. The findings suggest that ICT advances help the banking system perform better when transitioning from traditional to digital systems.

The following sections review previous literature about the ICT and its value and importance, methodology, result discussion, and conclusion.

2 Literature Review

The adoption of ICT has altered the content and quality of banking activities. According to all indications, ICT has much promise for re-engineering the banks' business processes. Investment in information and communication technology should be a key component of banking operators' overall strategy to maintain optimal performance. It is critical for bank executives to increase their investment in ICT goods to provide faster, more convenient, and accurate services, or they will lose ground to their competitors. The banking business provides a massive potential for ICT companies to sell their advances [9]. However, success in this field is contingent on their ability to adapt their offerings to appeal to the ready minds of various industry stakeholders.

The overhead cost reduction is the most significant advantage of ICT innovation in the commercial banks' operations and activities. Typically, maintaining physical branches, marketing, and labor are appreciably becoming more and more affordable

[10]. Kozak [11] investigates the ICT evolution's influence on the banking system's profit and cost-effectiveness over the stipulated period of 1992–2003 and finds a significant impact of the ICT execution on banking performance and cost-efficiency. Research by Bitler et al. [12] shows a significant difference in the business performance between banks with and without ICT innovation. In this line of research, Leckson-Leckey et al. [13] documented that ICT investment in higher-ICT-level banks has more advantage in improving banking profitability (ROA and ROE) than that in lower-ICT-level banks. Agu and Aguegboh [14] on 35 African banks in the 2013–2015 period by generalized method of moment (GMM) suggested that ICT had mainly affected short-term banking profitability. However, in the long term, ICT investment was making significant benefits for banking performance.

Meanwhile, some other studies do not support the overemphasis on technology investment when considering the theory of efficiency and cost. The acquisition of technology does not increase the efficiency of the banking business because of the weakness in operating technology and not optimally exploiting the benefits of technology for organizational purposes [15]. The study highlighted technology as not the critical factor. In order to increase business efficiency, banks should comprehensively combine the system of technical and social factors within the organization. Similarly, the combination of technology and human factors will make the organization more complex, challenging to form competitive advantage, and gradually lose the inimitable characteristics of the bank. As a result, bank profitability is significantly affected [16].

In general, the impact of ICT on bank performance is positive as ICT boosts productivity. However, the results may not be the same between different markets or regions. The banking system is the backbone of the economy, and hence, examining factors that affect bank performance is essential. To the best of our knowledge, the number of previous studies on the nexus between ICT development and bank performance is insufficient. This study aims to answer a research question: Is there a significant and positive nexus between ICT development and bank performance in Vietnam, a frontier bank-based economy.

3 Research Methodology and Result Discussion

3.1 Research Models

Based on the aforementioned theoretical study, ICT innovation offers commercial banks opportunities to improve their traditional business model more efficiently with lower operational costs and more qualified products. Moreover, ICT investment supports banks to enhance their risk control and management ability. In consequence, banks create more competitive advantages and attract more customers. The research model is presented as follows:

Table 1 Variable notes

Type variable	Name	Definition	Expected sign
Dependent variables	ROA	Profit after tax/total assets	
	ROE	Profit after tax/equity	
Independent variables	ICT index	Collected from the Ministry of Information and Communication	(+)
	SIZE	The log value of a bank's total assets	(+)
	ETA	Equity/total assets	(−)
	DLR	Customer's deposits/loans	(+)
	LLP	Provision of loss loans/total loans	(−)
	GDP	GDP growth	(+)
	INF	Inflation rate	(−)

$$\text{Performance}_{it} = \beta_0 + \beta_1 \text{ICT}_{it} + \beta_2 \text{Size}_{it} + \beta_3 \text{ETA}_{it} + \beta_4 \text{DLR}_{it} + \beta_5 \text{LLP}_{it} + \beta_6 \text{GDP}_{it} + \beta_7 \text{INF} + \varepsilon_{it}$$

In the regression equation, a commercial bank's performance is measured by ROE, ROA. The key explanatory variable, ICT Index, measures the readiness for IT development and application. The variables group of bank characteristics includes total asset (size), equity to assets (ETA), deposit loan ratio (DLR), and loan loss provision (LLP) [17]. Macroeconomic variables such as economic growth (GDP growth) and inflation (INF) play the role of additional controls [18]. Table 1 provides research model variable descriptions.

3.2 Data

Research data is collected from the annual financial statements of 39 commercial banks, accounting for approximately 90% of the market share of the banking industry in Vietnam from 2009 to 2020. ICT Index is obtained from the Ministry of Information and Communication data for commercial banks in the same period. A balanced panel data of 203 observations were synthesized, processed, and designed using Stata software. Descriptions of the data are shown in Table 2.

Table 2 Descriptive statistics of data

Variable	Obs	Mean	Std. Dev	Min	Max
ROA	203	0.00653	0.00133	0.00319	0.00965
ROE	203	0.07721	0.02355	0.03533	0.21043
ICT index	203	0.62940	0.06480	0.37280	0.81140
SIZE	203	19.00689	1.13980	16.20776	21.17315
ETA	203	0.08144	0.03338	0.02621	0.25539
DLR	203	1.14427	0.20032	0.71859	1.73678
LLP	203	0.01349	0.00540	0.00513	0.03979
GDP	203	0.02195	0.01278	0.00250	0.03540
INF	203	0.05994	0.01296	0.02910	0.07080

4 Results and Discussions

The correlation coefficient matrix between the variables (Table 3) shows the ICT Index variable has a strong positive correlation for ROA and ROE, showing a significant impact of information and communication technology investment on bank efficiency. The independent variables have a relatively low correlation with each other, so we can predict that there is no multicollinearity in the regression model.

Verification of multicollinearity phenomenon: The result is VIF < 3. Therefore, the variables in the model do not have the multicollinearity phenomenon (Table 4).

The research continues to perform a regression analysis to measure the trend and impact level of the independent variables on the dependent variables by methods such as Pooled OLS, FEM, and REM. Regression results according to Pooled OLS method with dependent variables ROA (1) and ROE (2), FEM method with variables

Table 3 Correlation coefficients between variables

	ROA	ROE	ICT index	SIZE	ETA	DLR	LLP	GDP	INF
ROA	1.0000								
ROE	0.7532	1.0000							
ICT index	0.4096	0.6181	1.0000						
SIZE	0.4599	0.5035	0.6989	1.0000					
ETA	−0.2435	−0.3085	−0.3731	−0.5778	1.0000				
DLR	0.0801	−0.0988	−0.2033	−0.1363	−0.0303	1.0000			
LLP	0.0207	0.0793	0.2242	0.3230	−0.1896	0.1049	1.0000		
GDP	0.1298	0.0413	−0.0215	0.1163	−0.2200	−0.1559	−0.2137	1.0000	
INF	−0.1946	−0.1564	−0.0571	−0.0723	−0.0540	0.0085	−0.0388	0.0337	1.0000

Table 4 Multicollinearity test

Variable	VIF	1/VIF
ICT index	2.74	0.365458
SIZE	2.06	0.484893
ETA	1.62	0.615889
DLR	1.22	0.816740
LLP	1.19	0.840061
GDP	1.12	0.894474
INF	1.02	0.980275
Mean VIF	1.57	

ROA (3) and ROE (4), REM method with variables ROA (5) and ROE (6), all of which are presented in Table 8.

F-test results of two models: Prob > $F = 0.0000 < \alpha$ ($\alpha = 1\%$): Hypothesis H0 is rejected: FEM will be more suitable than Pooled OLS.

Next, we perform the Hausman test for both models to obtain the following results (Table 5):

After finding a suitable method, REM, we test the model's defects such as variable variance and autocorrelation in Tables 6, 7 and 8 below.

Table 5 Hausman test result

H0: difference in coefficients not systematic			
Model (1): ROA		Model (2): ROE	
Chi2 (11)	11.26	Chi2 (7)	7.32
Prob > chi2	0.4568	Prob > chi2	0.3962

Hypothesis H0 is accepted: REM will be more suitable than FEM

Table 6 Result of LM—Breusch and pagan Lagrangian multiplier test

Model (1): ROA		Model (2): ROE	
chibar2(01)	69.90	chibar2(01)	422.68
Prob > chibar2	0.0000	Prob > chibar2	0.0000

Hypothesis H0 is rejected. Both models have variable variance phenomena

Table 7 Result of Wooldridge test for autocorrelation in panel data

H0: no first-order autocorrelation			
Model (1): ROA		Model (2): ROE	
F (1,25)	0.751	*F* (1,25)	1.028
Prob > *F*	0.3944	Prob > *F*	0.3204

Hypothesis H0 is accepted. Both models have no autocorrelation

Table 8 Results of regression analysis according to the models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ROA	ROE	ROA	ROE	ROA	ROE	ROA	ROE
ICT index	0.005***	0.194***	0.010***	0.159***	0.008***	0.166***	0.005***	0.104***
	(0.002)	(0.029)	(0.002)	(0.030)	(0.002)	(0.028)	(0.001)	(0.034)
SIZE	0.000***	0.003	0.002***	0.017***	0.001***	0.010***	0.001***	0.005***
	(0.000)	(0.002)	(0.000)	(0.003)	(0.000)	(0.002)	(0.000)	(0.001)
ETA	0.002	−0.042	−0.000	−0.004	−0.001	−0.022	0.006**	−0.010
	(0.003)	(0.049)	(0.003)	(0.039)	(0.003)	(0.038)	(0.002)	(0.029)
DLR	0.001***	0.004	0.002***	0.026***	0.002***	0.024***	0.002***	0.012***
	(0.000)	(0.007)	(0.000)	(0.006)	(0.000)	(0.006)	(0.000)	(0.004)
LLP	−0.037**	−0.421	−0.008	−0.221	−0.036**	−0.377*	−0.044**	−0.494***
	(0.016)	(0.264)	(0.018)	(0.226)	(0.017)	(0.219)	(0.013)	(0.183)
GDP	0.012*	0.028	−0.012	−0.084	0.014**	0.061	0.014***	0.057
	(0.007)	(0.110)	(0.008)	(0.099)	(0.006)	(0.085)	(0.005)	(0.071)
INF	−0.017**	−0.227**	−0.010**	−0.162**	−0.016**	−0.196***	−0.012**	−0.178***
	(0.006)	(0.100)	(0.005)	(0.059)	(0.005)	(0.060)	(0.004)	(0.055)
_cons	−0.006**	−0.076**	−0.032**	−0.362**	−0.010**	−0.233***	−0.010**	−0.088***
	(0.002)	(0.034)	(0.005)	(0.058)	(0.003)	(0.042)	(0.002)	(0.020)
N	203	203	203	203	203	203	203	203
adj. R^2	0.294	0.397	0.481	0.477	0.000	0.065	0.000	0.018

After fixing the defects of the two models, the results are presented in Table 6, with (7) being model 1 and (8) being model 2.

According to the REM, regression results after overcoming the defects of the two models: the regression coefficient of the variable ICT index with the two models 1 and 2 is 0.005 and 0.104, respectively, along with the 99% confidence level, showing that the level of readiness for IT development and application has a positive impact on the performance of banks.

5 Conclusions and Recommendations

In Vietnam, the legal mechanism for digital banking development still has many limits and constraints. In particular, there is no standard regulation and guideline for data connection and sharing among Vietnamese commercial banks. All current activities of connecting and sharing banking data are mainly bilateral with various forms of agreement, consuming more time and effort depending on different organizations. Besides, as the National Identity Database system for citizens is still developing, banks have many difficulties comparing and verifying customers' online (e-KYC). Therefore, the State Bank of Vietnam needs to develop laws to overcome the above shortcomings to support the Vietnamese banking system comprehensively approaching and developing ICT.

Along with the widespread application of technology, Vietnam's trend of high-tech financial and banking crime is on the rise both in sophistication and complexity, with many consequences leading to severe results. Currently, in the world, even developed countries with the most advanced technology have not found a perfect solution to ensure network safety and security. Therefore, the Vietnamese banking system needs to make efforts to update advanced security software suitable for the infrastructure, and at the same time, regularly warn customers about pretending to be a bank and appropriation of assets of the customer.

Investment in IT must be associated with training people to use in order to exploit technology effectively fully. To do this, banks should first standardize the recruitment conditions for entry. At the same time, it is recommended to organize classes to improve skills in using IT, especially the bank's IT department staff, because the current development speed in the world is rapid. This training must focus on operating and managing e-banking operations and responding and promptly handling customer incidents and vulnerabilities in technology. In addition, it is necessary to train professional sales staff. With the application of IT, more and more products and services are diverse in quality and geographical location. Therefore, bank staff needs to understand the product and online operation process fully and quickly process electronic documents to bring customers a sense of security, trust, and convenience.

Banks must focus on investing in IT infrastructure with their potential. The regular increase in total assets will help the bank strengthen its capacity, create a strong potential for further development of IT, modernize banking operations, help cut costs, and

help bank profits grow every day. In developing the digital economy, payment intermediaries are banks and FinTech companies, telecommunications, electronics, and information technology carriers. Therefore, the cooperation between the banking system and FinTech companies will bring high efficiency to the parties. This cooperation enriches products and services that are safe and convenient for users and solves the problem of excessive investment in the technology of the banking system. Social networks such as Facebook, Instagram, and Zalo are platforms with many visitors, a place to reach customers widely with low costs indirectly. Therefore, banks should promote propaganda through these platforms, encouraging customers to access more and more electronic products and services of the bank.

Acknowledgements We want to thank the Board of Editors and the two anonymous referees for the constructive comments and valuable suggestions. Special thanks to Ho Yen Nhi for excellent research assistance. Any remaining errors or shortcomings are the authors' responsibility. To conduct this research, financial support from the Van Lang University and University of Economics Ho Chi Minh City is acknowledged.

References

1. Vo XV, Nguyen HH, Pham KD (2016) Financial structure and economic growth: the case of Vietnam. Eurasian Bus Rev 6(2):141–154
2. Brynjolfsson E, Hitt L (1996) Paradox lost? Firm-level evidence on the returns to information systems spending. Manage Sci 42(4):541–558
3. Oliner SD, Sichel DE (2000) The resurgence of growth in the late 1990s: is information technology the story? J Econ Perspect 14(4):3–22
4. Stiroh KJ (2002) Are ICT spillovers driving the new economy? Rev Income Wealth 48(1):33–57
5. Rouse M (2005) ICT (information and communications technology, or technologies). Dostupno na: http://searchcio-midmarket.techtarget.com/definition/ICT (27.02. 2013)
6. Polasik M, Wisniewski TP (2009) Empirical analysis of internet banking adoption in Poland. Int J Bank Mark
7. Rahman MM, Hamid MK, Khan MAM (2015) Determinants of bank profitability: empirical evidence from Bangladesh. Int J Bus Manage 10(8):135
8. Del Gaudio BL, Porzio C, Sampagnaro G, Verdoliva V (2021) How do mobile, internet and ICT diffusion affect the banking industry? An empirical analysis. Eur Manag J 39(3):327–332
9. Alawode AJ, Kaka EU (2011) Information and communication technology (ICT) and banking industry. Mediterr J Soc Sci 2(4):71–71
10. Hernando I, Nieto MJ (2007) Is the Internet delivery channel changing banks' performance? The case of Spanish banks. J Bank Finance 31(4):1083–1099
11. Kozak S (2005) The role of information technology in the profit and cost efficiency improvements of the banking sector. J Acad Bus Econ 2(1):34–38
12. Bitler MP, Robb AM, Wolken JD (2001) Financial services used by small businesses: evidence from the 1998 survey of small business finances. Fed Res Bull 87:183
13. Leckson-Leckey GT, Osei KA, Harvey SK (2011) Investments in information technology (IT) and bank business performance in Ghana. Int J Econ Financ 3(2):133–142
14. Agu C, Aguegboh E (2020) ICT and Bank performance in Sub-Saharan Africa: a dynamic panel analysis
15. Dewett T, Jones GR (2001) The role of information technology in the organization: a review, model, and assessment. J Manag 27(3):313–346

16. Barney JB (2001) Resource-based theories of competitive advantage: a ten-year retrospective on the resource-based view. J Manag 27(6):643–650
17. Pham DK, Ngo VM, Nguyen HH, Le TLV (2021) Financial crisis and diversification strategies: the impact on bank risk and performance. Econ Bus Lett 10(3)
18. Duy PK, Nga TTT (2021) Risk-taking behavior and ownership concentration the case of Vietnamese Banks. Revista Geintec-Gestao Inovacao E Tecnologias 11(3):1155–1164

Technology Devices in Sport Education, Technology Applications with Wearable Sensor and Accelerometer in Physical Exercises—And Ho Chi Minh and Lenin Ideologies on Physical Exercise for the Youth

Nguyen Dinh Trung, Dinh Tran Ngoc Huy, Pham Thi Bich Thao, Trung-Hieu Le, and Dinh Thi Hien

Abstract Together with technology applications in sports such as wearable sensors and accelerometer and digital technology applying in coaching athletes, the young generation will be future owners of the nation, therefore they need to practice and do physical exercises to have a strong body and good mind and spirit. The objective of this paper is to find out what kinds of technical applications (such as wearable sensors) and how they can be applied in sports games as well as in physical exercises. In reality, up to now there are many researches and interests in the field of robotics interaction (this is new trend) with the meanings to recognize athletes or human activity, i.e., identifying type of activity of human, from the received signal stream (image, accelerometer, etc.). In this field there are at least two various methods to classify problem of activity recognition. And this is depending on whether machine learning models with recognition or the type of sensor is used to collect signals of activity. Study showed that to observe human or athletes activities to increase recognition efficiency, people are simultaneously observed by environmental cameras while still wearing accelerometers. Digital technology can also be applied for coaching

N. D. Trung
National Economics University (NEU), Hanoi, Vietnam
e-mail: trungnd@neu.edu.vn

D. T. N. Huy (✉)
Banking University HCMC, Ho Chi Minh city, Vietnam

International University of Japan, Niigata, Japan

P. T. B. Thao · D. T. Hien
Thai Nguyen University of Sciences, Thai Nguyen, Vietnam
e-mail: thaopt@tnus.edu.vn

D. T. Hien
e-mail: hiendt@tnus.edu.vn

T.-H. Le
Dai Nam Univeristy, Hanoi, Vietnam
e-mail: hieult@dainam.edu.vn

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_17

athletes. Moreover, the camera could be worn in several different locations on the body such as on the chest, wrists, and head. The combination of wearing multiple sensors has also been conducted for specific applications (care and monitoring of health progress and assessment of recovery) which can be used in sport games. Last but not least, regarding physical education for school youth, President Ho Chi Minh identified as an important part of the national education of an independent and democratic Vietnam: "An education will train children to become useful citizens for Vietnam, an education that fully develops their existing capacities."

Keywords Technology devices · Tech applications · Sport training · Wearable sensor · Accelerometer · Ho Chi Minh ideologies · Young generation · Vietnam

1 Overview

According to the sensor-based classification, there are two commonly used types of sensors: image sensors and accelerometers. Previous studies often used image sensors mounted in the environment to observe and acquire images of human activities. The use of sensors (cameras) mounted in the environment is often suitable for monitoring applications to monitor crowds in public places (building corridors, hospitals, airports, workplaces). This method does not require people to carry any equipment, but it is limited by the camera's field of view. The approach using accelerometer wear sensors is an option that allows to overcome this limitation. However, the signal received from the accelerometer is often quite ambiguous, making it difficult to identify and interpret the recognition results when necessary.

Under the socialist regime, science is the common property of the whole people. Therefore, the contingent of scientific and technical cadres must make every effort to spread their knowledge widely among the working people, so that the people can step up the emulation of producing more, quickly, well, and cheaply. Only then will the country be rich, the people will be strong, and the people's life will be improved in all aspects. That was President Ho Chi Minh's teaching to the national team of scientific and technical cadres exactly 57 years ago at the 1st National Congress of the Vietnam Association for the Dissemination of Science and Technology (Date May 18, 1963).

Uncle Ho's teachings emphasized the role and significance of science and technology in the cause of building socialism, so that the people could be rich, the country would be strong, and compete with the great powers of the five continents.

In summary, there is technology applications of wearable sensors for athlete coaches to record results of athletes in various sports: running, walking, jogging, mountain climbing, biking, etc.

On the other hand, digital technology also has applications in sports games as we presented in following section. And this sector received many financial support [1–9], and there are risks need to be managed [10, 11].

Research questions:

What are technology devices and applications in sports game and for athletes?

2 Research Approach and Methods

The article conducts an overview of the research works of domestic and foreign scholars, synthesize collected and searched documents to systematize theories related to technology applications, combined with survey, synthesis and inductive methods. Authors use observations and experiences, together with dialectical materialism, synthesis and inductive methods.

Santos et al. [12] mentioned effects of the skills4Genius sports-based training program in creative behavior. Then Stylianos et al. investigated activity recognition using wearable sensors for tracking the elderly. Beside, Tran et al. [13] stated a multimodal multi-view dataset for human fall analysis. Wearable Inertial sensors have revolutionized the way kinematics analysis is performed in sports. Therefore, this study differs from previous studies in a sense that it will identify supporting roles of digital tech for sports coaching and athletes. Also, athletes activities are recorded to increase recognition efficiency, they are simultaneously observed by environmental cameras while still wearing accelerometers. So there will come the novel of this paper.

In summary, Digital technology can be applied for coaching athletes.

3 Main Findings

3.1 Technology Devices and Applications in Sports

Haake [14] stated that with the help of simple equipment such as the pole vault or javelin considered as technological developments, which influence the index by around 30%, whereas one-hour record index with aerodynamic improvements is about 100%. Then, they come to conclusion that if distance or time is used as a performance measure, there is index of performance improvement extended to amateur and elite sport.

Next, Giblin et al. [15] mentioned there is a wide spread of applications of technology in many kinds of major sports and this will help to gain what they called competitive advantage in elite sports and this is important feature. There is innovation on information between athletes and their coaches, or method in which they collect and process data, so this has influenced greatly on how athletes are trained and monitored in training per day as well as competition environments.

Then, Kenttunen et al. [16] specified that there is increasing in demand of wellness technology for athletes for both the purpose of improving life and improving

Fig. 1 Volley ball as favorite sport. (*Source* authors collection from Vietnamese universities)

training quality, for instance, reducing injury risk. Besides, it is recognized that there is growing demand of information personalized going with the increasing sports interests (and wellness tech mentioned). There is also increase in digital coaching (with valuable training and guides) as it showed helpful for athletes' knowledge related to their techniques, for example in skiing technique perceived.

Figure 1 shows that volley ball as favorite sport which is collected from Vietnamese universities.

3.2 Technology Applications in Sports and Physical Exercises

3.2.1 Wearable Sensor and Accelerometer

With progress in technology, compared to whiteboards and reviews of post-practices, advanced technology has helped to increase potential of athletics as these tech became more resilient and smaller, more resilient, as well as less burdensome in recent years, For instance, wearable sensors can help to transfer to coach's tablet and GPS accurately pinpoints motion (with real time information conveyed), with support of smart phones and wearable tech help to prevent injuries. Compared to whiteboards and post-practice reviews, technology has substantially increased athletic potential.

Hence, training of sports are revolutionized with advanced technology through live-tracking performances, perfecting movements of athletic, as well as eliminating

injuries virtually. (source: https://onlinemasters.ohio.edu/blog/how-technology-is-revolutionizing-sports-training/, access date 9/21/2021).

In a few recent studies, to observe human activities to increase recognition efficiency, people are simultaneously observed by environmental cameras while still wearing accelerometers as studied by Tran et al. [13]. Even so, such a multimodal information acquisition system still faces difficulties such as installation cost, field of view and the difference between viewing angles. Recently, image sensors have been widely used to track the journey/daily activities of people in a number of sports and health monitoring applications. In these studies, the camera could be worn in several different locations on the body such as on the chest, wrists, and head. The combination of wearing multiple sensors has also been conducted by a number of research groups, but currently the methods are not geared toward user convenience (each sensor is worn in one position) so it is only suitable for specific applications (care and monitoring of health progress and assessment of recovery) that are not yet aimed at general users.

Wearable image sensor: In the past few years, there have been many studies using wearable image sensors to identify people's activities. Bearing image sensors generate large volumes of first-person image data, opening up new possibilities in gesture-based human–machine interaction (HCI) applications, activity logging, and gestures/sign language recognition. Usually, the camera is worn on the head or in front of the chest and neck to create a field of view similar to the human eye as shown in Fig. 2.

Fig. 2 Some common image sensor installation locations (*Source* Trung-Hieu Le thesis [17])

3.3 Ho Chi Minh and V. I Lenin Ideologies on Physical Exercises

"…Every weak citizen means the whole country is weak, every healthy citizen means the whole country is healthy.

… If the people are strong, the country is prosperous. I hope my compatriots all try to exercise. I practice every day by myself" are the sentences in Uncle Ho's "Call to exercise" posted on March 27, 1946. Talking about the goal, the people are strong and the country is prosperous are the two noble goals of the regime. This means that people's health is one of the great factors determining the development of the country toward "rich people and strong country". The people's health also contributes to the defense and construction of the country.

Therefore, other views of Ho Chi Minh are consistent with the orientation of strong development of sport for the health of the people. Through serving the people's health, serving the health of everyone, sport contributes to all activities of economic, cultural, social, and educational development… that is, to the cause of striving for "Rich people, strong country".

Ho Chi Minh's views on the development of mass sport President Ho Chi Minh pointed out that: "Under democracy, Sports and Physical Education must become common activities of the masses, aiming to enhance the people's health. If the people are healthy, everything can be done well." He encouraged: "So exercising, improving health is the duty of every patriotic citizen". Since then, Ho Chi Minh has advocated "We should develop the sports movement widely". Those are the general views of Ho Chi Minh about mass sport and sport for everyone. Regarding physical education for school youth, Ho Chi Minh identified this as an important part of the national education of an independent and democratic Vietnam: "An education will train children to be useful citizens for Vietnam, an education that fully develops their existing capacities". (*source: baoquangbinh.vn, access date 9/20/2021).*

V. I Lenin preferred to challenge himself (like many other sportsmen), which means physically, and he derived certain pleasure from being close touch with nature.

Elwood [18] stated during in Western Europe in his long years Lenin continued to pursue some sports, in that time he has become a mountain climber and a long distance cyclist.

In summary, the need of sports games and physical education are emphasized more by V. I Lenin and Ho Chi Minh and there are wide applications of devices including Wearable sensor and accelerometer in these sport activities.

4 Conclusions

In our modern society, science and technology can be applied in physical exercises and sport games as well as into other industries.

Technology applications and devices and wearable sensors have been applied into many sport games and has helped many athletes to reduce injuries (so, the camera could be worn in several different locations on the body such as on the chest, wrists, and head).

Beside, athletes activities are recorded to increase recognition efficiency, they are simultaneously observed by environmental cameras while still wearing accelerometers. Digital technology can also be applied for coaching athletes.

Diehl et al. [19] said physical activity among students is essential for complimenting sedentary behavior and for individuals' future health.

President Ho Chi Minh attaches great importance to physical education for the young generation, because the young generation will be future owner of the nation. On March 31, 1960, President Ho Chi Minh wrote "Letter to the Conference of Sports Officers in the North", the teacher: "If you want to have good productive labor, good work and study, you need strength strong. If you want to stay healthy, you should develop a widespread sport movement". The person advised: "Sports and physical training officials must study politics, conduct professional research and work enthusiastically to serve the people's health." (source: Ho Chi Minh: Complete Volume, Publishing House. National politics, Hanoi, vol. 12, p. 542).

Fig. 3 Ho Chi Minh paid attention to physical exercise. (*Source* authors collection from tuyengiao.vn)

Figure 3 shows that how Ho Chi Minh paid attention to physical exercise which is collected from tuyengiao.vn.

Research Limitation

We need to make detailed analysis for each sport games training.

Acknowledgements Thank you editors, friends and brothers to assist this publishing. Warm thanks to Mr. Trung-Hieu Le who supports some figures.

References

1. Huy DTN, Linh TTN, Dung NT, Thuy PT, Van Thanh T, Hoang NT (2021) Investment attraction for digital economy, digital technology sector in digital transformation era from ODA investment-and comparison to FDI investment in Vietnam. Laplage em Revista 7(3A):427–439
2. Huy DTN, Le TH, Thang TD, Hoa NT, Hue LT (2021) Discussion on E-learning solutions for students–and issues of technology application in classroom. Des Eng (Toronto) 7:11432–11443
3. Hac LD, Huy DTN, Thach NN, Chuyen BM, Nhung PTH, Thang TD, Anh TT (2021) Enhancing risk management culture for sustainable growth of Asia commercial bank—ACB in Vietnam under mixed effects of macro factors. Entrepreneurship Sustain Issues 8(3):291–307
4. Hanh HT, Chuyen BM, Huy DTN, Phuong NTT, Phuong HTL, Nuong LN (2020) Challenges and opportunities from impacts of FDI and income on environment pollution: role of financial accounting transparency in FDI firms. J Secur Sustain Issues 10(20):643–654. https://doi.org/10.9770/jssi.2020.10.2(22)
5. Huy DTN, Loan BT, Anh PT (2020) Impact of selected factors on stock price: a case study of Vietcombank in Vietnam. Entrepreneurship Sustain Issues 7(4):2715–2730. https://doi.org/10.9770/jesi.2020.7.4(10)
6. Huy DTN, Dat PM, và Anh PT (2020) Building and econometric model of selected factors'impact on stock price: a case study. J Secur Sustain Issues 9(M):77–93. https://doi.org/10.9770/jssi.2020.9.M(7)
7. Hang NT (2017) Testing the influence of factors on the industrial development of Thai Nguyen province. J Ind Trade 7:82–89
8. Hang NT (2016) Impact of foreign direct investment on the development of industry in Thai Nguyen province. J Econ Manage-Ministry Plann Investment 7+8(77):70–80
9. Nghia LH, Van Chien N (2014) Impact of foreign direct investment on labor productivity and technology level of Vietnam. National Political Publishing House, Hanoi
10. Huy DTN (2015) The critical analysis of limited South Asian Corporate Governance Standards after Financial Crisis. Int J Qual Res 9(4):741–764
11. Huy DTN (2012) Estimating Beta of Viet Nam listed construction companies groups during the crisis. J Integr Dev 15(1):57–71
12. Santos S, Jimenez S, Sampaio J, Leite N (2017) Effects of the skills4Genius Sports-based training program in creative behavior. Plos One 2
13. Tran TH et al (2018) A multi-modal multi-view dataset for human fall analysis and preliminary investigation on modality. In: 24th international conference on pattern recognition (ICPR). IEEE
14. Haake SJ (2009) The impact of technology on sporting performance in Olympic sports. J Sports Sci 27(13). https://doi.org/10.1080/02640410903062019
15. Giblin G, Tor E, Parrington L (2016) The impact of technology on elite sports performance. Sensoria 12(2). https://doi.org/10.7790/sa.v12i2.436

16. Kenttunen E, Kari T, Makkonen M, Critchley W (2018) Digital coaching and athlete's self-efficacy—a quantitative study on sport and wellness technology. In: The 12th mediterranean conference on information systems (MCIS), Corfu, Greece
17. Le TH, Huy DTN, Hang NT, Dung NT, Thanh NTP, Gwoździewicz S, Vu H (2021) Human machine interaction and meanings of machine learning-a case of hand posture recognition from wrist-worn camera. Des Eng (Toronto) 7:11174–11187
18. Elwood C (2010) The sporting life of V. I. Lenin, Canadian Slavonic Papers. Revue Canadienne des Slavistes 52(1/2):79–94
19. Diehl K, Fuch AK, Rathmann K, Kolb JH (2018) Students' motivation for sport activity and participation in university sports: a mixed-methods study. Biomed Res Int 3https://doi.org/10.1155/2018/9524861

Applications of Statistic Softwares, IT and Artificial Intelligence—A Research Case on ByPass Behavior of Patients at Vietnam Hospitals

Pham Van Tuan and Dinh Tran Ngoc Huy

Abstract Nowadays, in industry 4.0, Artificial intelligence, information technology and statistic softwares such as AMOS and SPSS has influenced the researches conducted in medical sector and hospitals. The increasing number of patients by pass at hospitals in Vietnam caused by their self-transfer recently which increase costs (financial) and time and cross-infection risk at hospitals. From November 2019 to February 2020, authors conduct a Quantitative research with a survey sample of 315 patients in the Northern provinces of Vietnam. Next authors have used software (IT statistics software), in our survey taking place at populous provinces, such as AMOS 22.0 for analyzing CFA affirmation component and then analyzing EFA discovery component with SPSS (also IT software), based on interviews conducted for medical doctors and expert group and data of patients features and their reasons to overtaking. Our research findings show that among factors that impact (positively) on the intention to bypass to higher-level hospital including Reference group (less impact than others 0.249 and influence positively on attitude 0.177), Awareness, and Attitude. We will explore gap of research in a sense that effect of the reference group on attitudes and behaviors bypassing of patients in Northern Vietnam.

Keywords Statistic softwares · AMOS · Applications · Reference group · Attitudes · Behaviors · Hospital self-transferring

1 Introduction

In this study, we will use information technology and statistic softwares such as AMOS 22.0 and SPSS in order to address one of the urgent medical problems in recent

P. Van Tuan (✉)
National Economics University, Hanoi, Vietnam
e-mail: Phamvantuan@neu.edu.vn

D. T. N. Huy
Banking University HCMC, Ho Chi Minh City, Vietnam

International University of Japan, Niigata, Japan

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_18

years in Vietnam is hospital overcrowding due to a large number of patient bypassing to high-tier hospitals. Bypassing patients, or patients' hospital self-transferring, is a behavior in which patients receive health care from more distant providers than the health care provider closest to their place of residence. In 2018, according to the practical data from Ministry of Health more than 35% of patients come to the central level (Level I) for treatment but can be treated at the provincial level (Level II). At the provincial level, up to 41.5% of patients could be treated at the district level (Level III) and 11% could be treated at the commune health station (Level IV).

Comparing to other statistics software such as R, Stata, or SPSS AMOS (which help us to perform structural equation modeling (SEM) to build models with more accuracy) advantages of SPSS include first, applications for investigation of sociological and econometric fields; second, its friendly user-interface, and third, non-parametric tests (Chi-square, Phi, lambda…) performance, and fourth, power in analyzing linear regression simple and multiple, mean test (T-test),descriptive statistics, "reliability testing of the scale" or, cognitive mapping (used in marketing) or use using price variables, etc. This cannot be done if only using SPSS software in case the model has many dependent and intermediate variables. With SPSS Amos, you can define, estimate, evaluate, and present your model in an intuitive interface that shows the relationships between hypothesized variables. SPSS Amos also provides a non-graphical method for model definition. SPSS Amos is the perfect tool for many purposes, in addition to allowing VB# or C# programming to be used to design models. The operating system supports Windows. The data format for Amos usually uses the input file format of SPSS. CFA analysis, SEM test, bootstrap analysis, multi-group analysis, impact analysis of regulatory variables, intermediate variables were performed with this AMOS software.

The more the local economy develop and grow, the more patients and overload at hospitals.

Moreover, the lower-tier hospitals have not fulfilled the task as "the gate guard" which monitor chronic non-communicable diseases, manage primary health care, and educate people about health prevention and promotion [1]. Doctors also do not have many conditions to scrutinize difficult cases so their medical professionalism might fade away. Also, the weakness of the lower-level hospital makes people only go to see the doctor until their illness gets worse, and then they will go straight to the provincial and central levels.

Previous studies have not fully examined factors affecting the Vietnamese patient's intention to bypass. In addition, little is known about the application of the TPB behavioral intent theory to the issue of hospital overcrowding in Vietnam. Our research goal was to test elements associated with the choice of hospital bypass from lower to higher-level hospitals. Findings from the study can help the government improve and utilize lower-tier hospitals as well as implement suitable health strategies.

2 Theoretical Framework

2.1 Relationship Between Reference Group and Perceived Behavioral Controls

Sherif's Autokinetic Effect Experiments (1935) proves that people's perceptions are affected by the community. From the research, Sherif found that in a situation which things are ambiguous, people will adapt to the group's denominator, which is a testament to the informational influence. In addition, Levy's research, Lee describes the decisions of family members have a strong influence on the individual.

We have several hypotheses as follows:

Name	Content	Description
H1	*The reference group effects the perceived behavioral control*	Our attitudes and behaviors can be strongly shaped and often inadvertently shaped by those around us
H2	*The strong effect of the reference groups toward the attitude*	Adams et al. (2008) [2] found that hospital characteristics such as distance between hospital and place of residence are the negative factors to patient's intention to self-transfer to higher-hospitals, beds positively impacting bypassing intention
H3	*The elements of hospital also influence the intention of bypass to the higher-hospitals*	In the research from Werner de Cruppé et al. most patients choose to be treated at the health facilities which they know the most, meaning they have lots of perceived behavioral control about them. These include perceived behavioral control of past experience (more than 50%), perceived behavioral control of the advice of the reference group, and awareness of the distance between the treatment site and their home
H4	*The effect of perceived behavioral control is strong*	With 4,232 outpatients referred to examination, Birk et al. (2011) [3] found that 22% of outpatients consider their past personal experience influenced their hospital choice in the future, while 7% were influenced by your experience. Friends and 5% by family experience. Only 3% refer to the media

(continued)

(continued)

Name	Content	Description
H5	*The reference groups persuade patients to self-transfer to a higher-hospital*	Attitude will reflects personal positive or negative assessment on behavior performing, and reflect his or her feeling about target behavior. Ajzen [4] specified that if individual has positive assessment of behavior performing, the intention to perform a behavior will be higher
H6	*The intention of bypassing can be created by the affection of attitude*	

2.2 Relationship Between Peripheral Factors (Ngoại Biến) and Self-transfer to Higher-Hospitals

2.2.1 Demography

Buczko (1992,1994) [5, 6] surveyed 670 patients who used medical services in Delaware in 1987 and found that age, sex, length of stay and severity of the disease are factors that influence the health facilities selection. Goldsteen et al. used the discharge data of 2171 inpatient in rural Illinois who found that age, payment source, and per capita income are significant determinants of hospital options.

2.2.2 Geography: Vietnamese North Region

The North has complicated terrain, and building the large hospitals in the mountains is difficult to constructed. According to General Statistics Office Of Vietnam (2018) [7], if counting the North alone, the Red River Delta has 74,014 beds, while the Northern Midlands and Mountains region has only 38,755. However, the Government of Vietnam has come up with many solutions to bring medicine closer to the people, namely, building a tighter health network and stronger grassroots health system. In 2017, the Ministry of Health surveyed and selected 26 health stations in 8 provinces and cities, according to 3 regions to deploy and timely adjust in the replication process. The commune health stations follow the Instruction No. 1383/HD-BYT [4] on December 19, 2017 of the Ministry of Health. In particular, mainly renovating the operation of the Health Station according to the principle of family medicine. This can help reduce the hospital overcrowding when the number of beds is not evenly distributed.

In addition, higher-level hospitals can open the second facilities or extend the size so that they can help patients avoid risks during treatment. The rational distribution

of the health facilities network is a factor that influences the intention to self–transfer of the people in general and the North of Vietnam in particular.

2.2.3 Policy

The way of classification and ranking of hospitals is still inadequate and has not created an appropriate motivation for hospitals to improve the quality of medical examination and treatment. It can be seen that the current ranking of hospitals gives patients absolute confidence in high-class hospitals and ignores lower-level hospitals in local. This makes the upper level always under pressure despite the facilities and medical capacity of the lower-level is improving.

According to Clause 3, Article 22 of the Law on Health Insurance 2008 (amended and supplemented in 2014) [8], cases of medical examination and treatment not in line with the right conditions will have little or no benefit from insurance. However, with the current situation in Vietnam, patients still need to seek medical treatment at hospitals at higher levels.

Then, Authors use framework TRA and in the context of slicing variables such as geography, patients, and policies. This model has variables that are considered in different impacts:

3 Research Methods

Authors use preliminary research and formal research. In details, the official research questionnaire was put into a quantitative survey in Northern Vietnam. The survey sample includes patients who have been and will be intentionally crossing the line. From cities and provinces, the author chose an appropriate research sample. The author conducted patient interviews in Northern Vietnam according to a reasonable proportion of the population. Specifically, respondents in the profile were as follows: women accounted for 61.59% of men and women; 4 age groups (under 18, 18–30, 31–50, over 50). In our research, authors collected samples on a scale of 321 respondents ($n = 321$) and actual answer was 315 (questionnaire). After screening invalid questionnaires then removing, we used 315 valid responses for formal analysis. Then authors used SEM for testing scale and research model, the (Model equation structure). Detailed data checking is performed through the following steps: first step is evaluating (preliminary) scale and variables reliability with analysis of EFA discovery factor and Cronbach's Alpha coefficient; then second step is examining the uniqueness, aggregate reliability and value, with support of SPSS 22.0, AMOS software in SEM analysis and using confirmatory factor analysis (CFA).

The sample size includes 315 patients, of which 61.59% are female and 38.41% are male. The age of these subjects is mainly in the age group of 18–31 (40.31%), 31–50 (36.51%) and >50 (19.05%), the rest are aged <18 (4.13%). The patients using medical services, accounting for 46.35%, surgical services, accounting for 32.06%,

obstetric services accounting for 6.03%, pediatric services accounted for 3.49% and other services accounted for 12.06%. The current living regions of respondents are mainly highland (53.02%) and countryside (26.03%), the remaining is city under the province (19.68%) and central cities (1.27%).

4 Research Results

4.1 Artificial Intelligence and Information Technology Applying in Hospitals

The overload at hospitals might come from many reasons: the economic growth, and the shortage of hospitals and doctors and nurses, as well as imbalance between medical services in rural and urban regions.

Artificial intelligence, the English term is Artificial intelligence—AI, is a rapidly growing computer technology that has begun to be widely used in the medical field to improve professional qualifications and efficiency in clinical work. In developing countries around the world, there are still inequalities between urban and rural health services, of which the shortage of doctors is the main cause.

Nowadays, because of big data and information in medical services, doctors may need AI to help them to assess patients' documents, profiles, and support making treatment decisions.

The prospects of AI technology can help healthcare providers collect, store, reformat, and track clinical data, as well as personalized plans and assessments.

AI technology has been used to improve diagnostic quality, especially in radiology. AI based on data source of clinical images to diagnose skin diseases, the results proved that this system can classify skin cancer at the same level as dermatologists.

SPSS is widely used by researchers for research in the following areas:

- Psychology, criminology
- Sociological investigation: Assess the quality of public services, identify factors affecting people's perception, etc.
- Business research: Intention to buy a product, tendency to accept a product or service. Brand positioning on product and service attributes
- Research in biomedicine: Effects of drugs on a group of diseases, analysis of factors affecting agricultural pests…

SPSS, together with AMOS, also allows researchers to use high-order quantitative analysis techniques such as structural equation modeling (SEM) analysis to measure and test many theoretical models. Customer satisfaction: Use SERVQUAL models. Gronroos (Nordic Model), ROPMIS, etc. and variations from these models.

Acceptance Assessment of Information Systems: Using Technology Acceptance Model (TAM), Information Systems Success (ISS), TRA Model, TPB (Ajzen), E-Cam, etc.—Trend Assessment transformation direction, conversion barriers in the field of e-commerce and telecommunications.

4.2 Positive Factor Analysis (CFA) Testing Models

4.2.1 Compare Market Segments by Demographic Criteria

The NPar test shows that the age of patients who have the highest level of hospital bypass is between 31 and 50 years old (186.48 [MOU1]).

Similarly, when applying the NPar test to geographic areas, including mountainous, rural, provincial cities, and central cities. The results show that patients in rural areas have the highest intention to bypass local hospital (199.32).

	Age	*N*	Mean rank		Area of resident	*N*	Mean rank
tbINT	Under 18	13	135.96	bINT	Highland	167	146.38
	18–30	127	156.16		Rural area	82	199.32
	31–50	115	186.48		Provincial cities	62	136.63
	Over 50	60	112.08		Central cities	4	127.38
	Total	315			Total	315	

In addition, the research team continued to use the NPar test with the medical examination and treatment services used by the patient when bypassing hospitals. Among the internal medicine, surgery, obstetrics, pediatrics, and other services (geriatric, ophthalmic, psychiatric, infectious …) patients have the most desire to bypass when using the internal medicine service (171.64). In addition, patients who want to use surgical services also have much higher bypass intention (164.20) than patients who need to use obstetrics, pediatrics, and other services.

	Medical services used when bypass	*N*	Mean rank
tbINT	Medical	146	171.64
	Surgery	101	164.20
	Obstetric	19	116.71
	Pediatrics	11	122.14
	Other	38	120.16
	Total	315	

5 Discussion

5.1 Research Results Discussion

The research results and discussion has proved the intention to cross the line is influenced by many factors, both objective and subjective. To minimize the situation of bypassing, it is necessary to call the actions of departments, unions, and stakeholders.

This study mainly use statistics softwares such as AMOS 22.0 and SPSS and address roles of information technology and artificial intelligence in hospitals and medical sector.

First of all, the government needs to have policies related to health insurance, the budget for all levels of health and the need for more strict regulations on medical examination and treatment for out-of-line to manage over-route in the health system. In parallel with the above measures, the government needs to monitor the level of service improvement in urban and rural hospitals, especially grassroots health care, and primary health care. Regarding the quality of examination at lower-level hospitals, it is necessary to have policies to attract talents, or regularly organize seminars to improve the skills of physicians.

5.2 Limitations and Future Research Recommendations

Firstly, the authors' research sample is concentrated in Northern highlands, Nam Dinh province and scattered in some other Northern cities and provinces. In the meantime, due to the COVID-19 pandemic during the survey period, the number of samples was still modest. Although these are densely populated provinces and the sample size is considerably larger than the minimum, it is difficult to confirm whether the results would represent for people in other areas or not.

Secondly, although the article has attempted to concretize and express the research scales clearly, it might have led to difficulties in understanding and incorrect responses due to the limited medical knowledge of respondents. Future research studies would examine other factors affecting patients' behaviors and find suitable implement to reduce bypassing rate.

Acknowledgements This research is funded by National Economics University, Hanoi Vietnam.

References

1. Hac LD, Huy DTN, Thach NN, Nhung PTH,Thang TD, Anh TT (2021) Enhancing riskmanagement culture for sustainablegrowth of Asia commercial bank—ACBinVietnam under mixed effects of macrofactors. Entrepreneurship Sustain Issues 8(3)

2. Tuan NTH et al (2019) The situation of bypassing of emergency and intensive care department at Hoc Mon regional Hospital in 2018–2019. Tuan NTH et al (2019) Status of referrals of emergency resuscitation department of Hoc Mon area general hospital in 2018–2019
3. Tran TN (2016) Thematic internship report: "Study on the influence of reference group towards the behavior to choose medical and examination service of North citizen, Viet Nam". Trần TN (2016) Chuyên đề thực tập "Nghiên cứu ảnh hưởng của nhóm tham khảo đến hành vi lựa chọn dịch vụ khám chữa bệnh của cư dân tại miền Bắc"
4. Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179–211
5. General Statistics Office of Vietnam (2018) Number of patient beds in 2018 by province
6. Health Strategy and Policy Institute (2013) Study on current situation of overcrowding, under-crowding in hospitals at levels and recommended solutions for improvement
7. Ministry of Health (2018) Health Report on October 12, 2018 Ministry of Health (2018) Health news on 12/10/2018
8. Nhat HB (2014) The inpatient satisfaction of medical and examination service in hospitals in Long Xuyen city. Nhật HB (2014). Sự hài lòng của bệnh nhân nội trú đối với chất lượng dịch vụ khám chữa bệnh của các bệnh viện tại thành phố Long Xuyên

Applications of IOTS, Internet Data, and Artificial Intelligence in Building Better Management Information System (MIS) in Banking Activities in Vietnam

Nguyen Van Thuy

Abstract In the context of digital transformation, first, the author emphasized the roles of Internet of Things and reliable Internet data which have become valuable for banking sustainability via building suitable regression model with support of statistics softwares such as Eview and SPSS serving for better management information system of banks. Without Internet technology, banks are harder to win more market share and expand marketing strategies. Using IT software (Eview), it uses practical bank cases: big listed banks in our country and take advantage of synthesis and inductive qualitative methods for analyzing. Besides, this study also presents concepts of applications of digital technology to organize online database for banks and cybersecurity for online bank database. Then, it will give out risks recommendations during Industry 4.0. Lastly, we suggest enhancing better MIS-management information system for improving bank sustainability in Vietnam economy. We will use IT softwares such as Eview, SPSS, or Stata in order to analyze risk factors step by step.

Keywords IoTs · Internet data · MIS · AI · Digital transformation · Vietnam

JEL Classification M21 · G30 · G32 · G38

1 Introduction

First of all, we recognize that Vietnam banks can take advantage of Internet technology together with Internet data in order to contribute for better bank MIS system.

Anh et al. [1] stated Vietnam banking sector has been developing fast over years and need better MIS system for risk data, and this is also confirmed by Thach et al. [2]; Huy et al. [3], Ngu et al. [4], Hue et al. [5], Anh et al. [1].

N. Van Thuy (✉)
Banking Academy, Hanoi, Vietnam
e-mail: thuynv@hvnh.edu.vn

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_19

In this paper, we mainly focus on taking Internet data that is reliable with IT software support (EViews) in comparing and evaluating macro-factors that affect market risk of Vietnam banks.

2 Literature Review

2.1 Previous Studies

Applications of digital technology to organize online database for banks:

Other modern technology such as cloud technology, storing data on cloud and Internet of Things (IoT) and AI has transformed banks activities into more digital approach.

Huy et al. also stated Vietnam banks need better risk models for risk MIS. Eugene [8] stated CAPM model related to risks.

And Huy [3] also specified that better governance in banking need better MIS and risk models. Dinh et al. [6] stated FDI investment also function and has vital role.

Next, we summarize as Table 1.

3 Methodology

Description.

We use EView software to make OLS regression.

1st step we use EView to make scatter charts.

2nd step we use EView to run OLS regression.

3rd step we discuss results.

4th step we generate policy implications.

Data we use from online Web sites that reliable such as SBV—State Bank of Vietnam, Bureau of Statistics, and bank sector.

4 Main Findings

4.1 Overview of Results

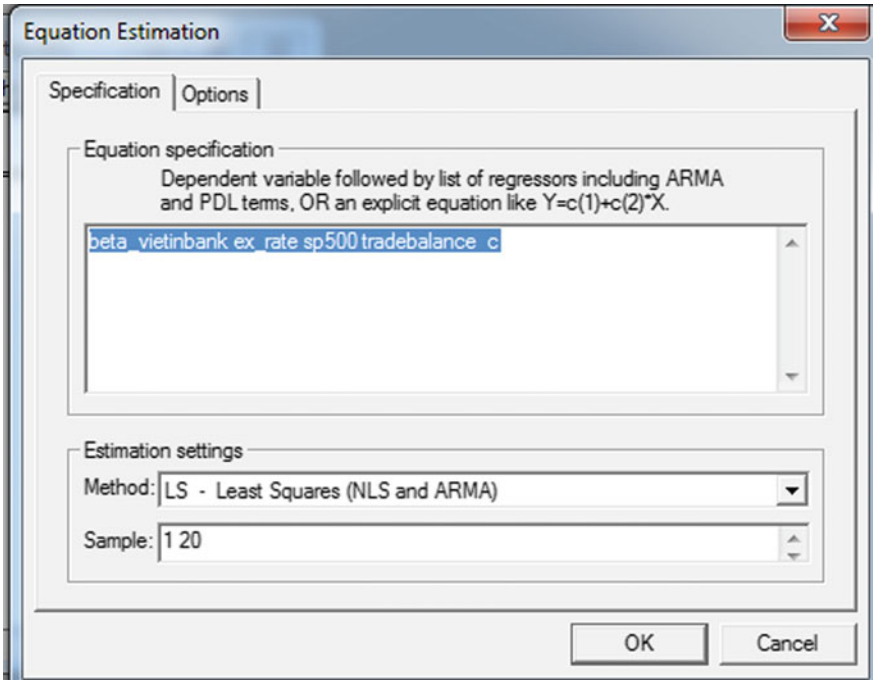
We look at below figures a examples of using EView to make scatter charts of correlation between macro-factors and stock price of Vietnam banks:

Next, we run OLS regression with IT software EView as follows:

Step 1: **Use EView software to run equation (OLS)—case of macro-elements**

Table 1 Summary of previous studies

Authors	Year	Contents, results
Avegrou, C	2008	We can improve information system or IS with social relations transformation
Dinh Tran Ngoc Huy, Dinh Tran Ngoc Hien	2010	Risk management is a vital component of corporate governance
Huy, Dinh T.N	2012	Risk also appear in construction sector
Gunaratha	2016	Note that risk (financial) is affected negatively by firm size and positively by leverage
Hami	2017	In survey and chosen period, inflation affects negatively on financial depth in case of Iran
Giebe et al.	2019	In bank sector, we need take advantage of "Big Data & Analytics" for improving customer-oriented services
Feitosa et al.	2019	Using disruptive technologies can support change in labor skills, relations, and communication
Kantos and Batolomeo	2020	Market index benchmark can be supported with "alpha" superior to smart beta
Dinh Tran Ngoc Huy, Nguyen Thi Thuy, Le Thi My Phuong, Pham Minh Dat, Vu Trung Dung, Phan Tien Manh	2020	Risk management need to be improved in listed firms
Pham Minh Dat, Nguyen Duy Mau, Bui Thi Thu Loan, Dinh Tran Ngoc Huy	2020	Risk governance is vital to firms
Dinh Tran Ngoc Huy, Pham Ngoc Van, Nguyen Thi Thu Ha	2021	Better research skills need to be trained in universities
Huy, D.T.N., Trung-Hieu, L., Hang, N.T., Sylwia, G., Trung, N.D., & Tuan, P.V	2021	Technology also may cause risks
Fan Yong-Yan, Jalil Manafian, Syed Maqsood Zia, Dinh Tran Ngoc Huy, Trung-Hieu Le, 2021; Nguyen Thi Huong Lan, Dinh Tran Ngoc Huy, 2021	2021	Mathematical model can be applied in suitable situation
Huong, D.T et al.	2021	Economic case can be taught in classroom



(source: author generate from Eview)

Step 2:

Analysis: We recognize from above tables (Table 2) that Ex_rate and tradebalance have positive effects on market risk, whereas SP500 has negative effect on beta of bank.

Table 2 Obtain OLS results

Variable	Co-efficient	Std. Error
Ex_rate	0.0001	0.0002
SP500	−0.0002	0.0003
Tradebalance	0.0003	0.0004
C	−2.2	5.7
R-squared	0.04	
SER	0.67	
Akaike info criteria	2.2	

source Author run from Eview

5 Discussion and Conclusion

In case of Vietinbank (previously SOE): For external risk factors, trade balance has more positive effect on market risk, compared to exchange rate. This can lead to policy implications.

In case of Vietcombank (VCB): From scatter charts above, we see elements such as IM, Rf has negative effect on beta (market risk) as well as negative relation with its stock price (see Charts 1 and 2).

These risks information and risk analysis can be produced under a risk report so that banks can deliver to investors via suitable MIS with access available online for suitable shareholders and investors.

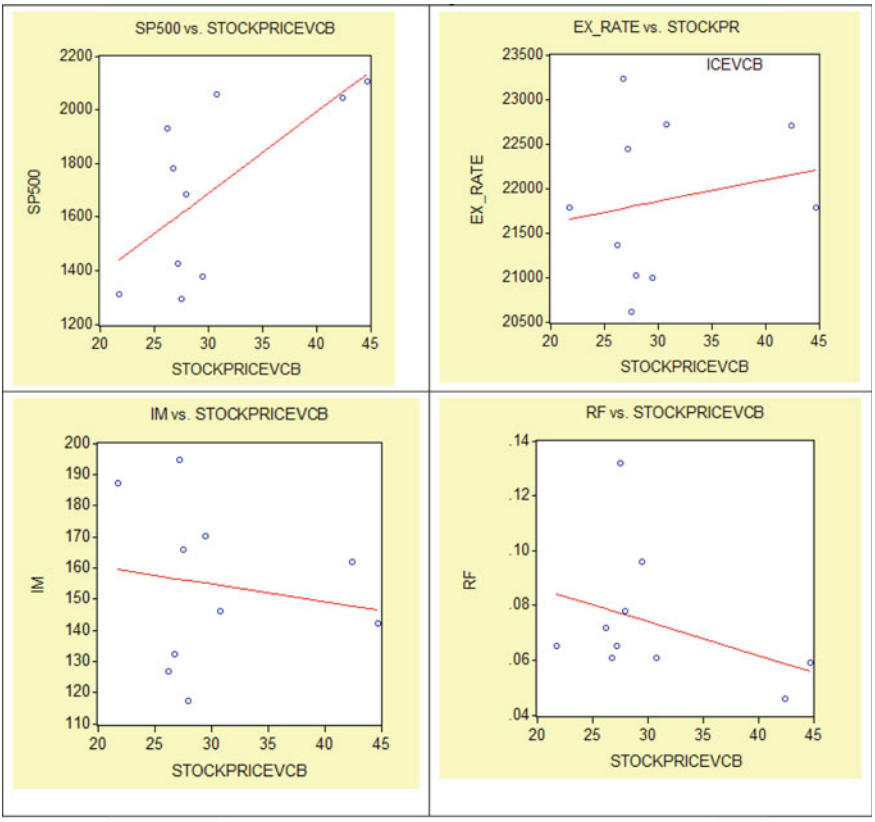


Chart 1 Stock price of VCB versus exchange rate, SP500

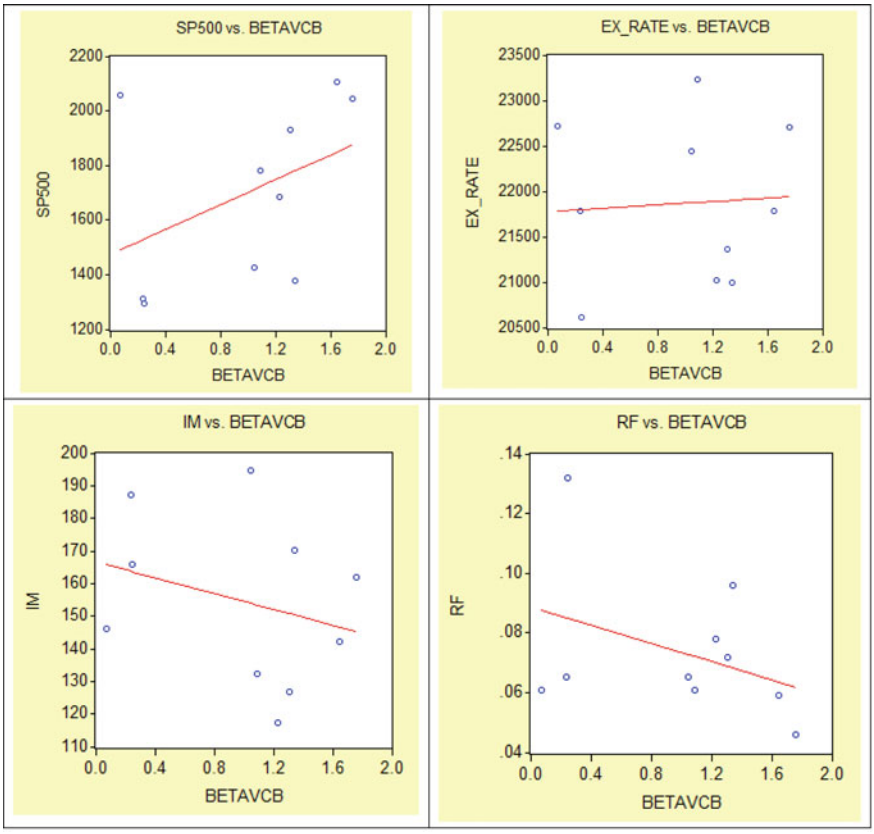


Chart 2 Beta CAPM of VCB versus exchange rate, SP500

6 Recommendations for Better MIS

We need to organize better online data for MIS, esp. We need to focus on cyberthreats and implement practices to defend threats and risks for computer, software, and hardware and avoid hackers to access our database.

Even our network and database in the banks also need to be protected against hackers. In recent years, there are some database stolen by hackers and harmful for banks and its clients. Thats why this study will suggest cybersecurity solutions.

Acknowledgments Authors send warm thanks to friends, editors, and co-workers.

References

1. Anh P, Huy DTN, Phuc DM (2021) Enhancing database strategies for management information system (MIS) and bank sustainability under macro effects—a case study in vietnam listed banks. Acad Strat Manage J 20(6)
2. Thach NN, Hanh HT, Huy DTN, Huong LTT, Nam VQ (2021) technology quality management of the industry 4.0 and cybersecurity risk management on current banking activities in emerging markets - the case in Vietnam. Int J Qual Res 15(3):845–856
3. Huy DTN, Trung-Hieu L, Hang NT, Sylwia G, Trung ND, Tuan PV (2021) Further researches and discussion on machine learning meanings-and methods of classifying and recognizing users gender on internet. Adv Mech 9(3)
4. Ngu DT, Huong DT, Ngoc Huy DT, Thanh PT, Döngül ES (2021) Language teaching application to English students at master's grade levels on history and macroeconomic-banking management courses in universities and colleges. J Lang Ling Stud 17(3):1457–1468
5. Lan NTH, Huy DTN (2021) Developing students' mathematical competence through equipping them with necessary knowledge about metacognition-and activities in teaching mathematics in secondary school. Laplage em Revista 7(3B)
6. Eugene FF, French KR (2004) The capital asset pricing model: theory and evidence. J Econom Perspect.
7. Dinh TNH, Tran TNL, Phan TT, Ta VT, Nguyen TH (2021) Investment attraction for digital economy, Digital technology sector in digital transformation era from oda investment and comparison to fdi investment in Vietnam. Laplage em Revista (International) 7(3A)
8. Huy DTN, Van PN, Ha NTT (2021) Education and computer skill enhancing for Vietnam laborers under industry 4.0 and evfta agreement. Elementary Educ Online 20(4)

COVID-19 Regulation Analysis Using Deep Learning

P. Baskaran, R. Rengarajan, S. Naveen Kumar, and V. Vijay

Abstract Due to ongoing worldwide pandemic, an approach for detecting social distance among neighbors is essential one. So, this work makes use of deep learning to calculate the position between individuals with the goal of reducing adverse impacts of the coronavirus epidemic. By analyzing a camera feed or video, we created a detection tool to alert people to keep a safe distance from each other. The camera's live footage or video was fed into an open-source object detection algorithm, which converted the video into a two-dimensional top view and calculated the distance between the people, giving us a rough estimate of their distance. The developed technique has the potential to be used in real-time implementation. Social distancing has proven to be a highly effective strategy for slowing the spread of virus. The main objective of this approach is to introduce automatic monitoring of social distance among human people and to alert them whenever there is violation of social distance among crowded people. This tool can be integrated into any security camera system to guarantee that the social distancing protocol is followed and tracked in civic places and workplaces.

Keywords Deep learning · Social distancing · Coronavirus · Pandemic · Safe distance · Alert

1 Introduction

To prevent the spread of a contagious virus is to practice social distancing. For several generations, it has been known that keeping a safe distance from an infected patient can help prevent the spread of an infectious, spreading virus. By delaying the spreading of virus, the probability of spreading the virus, hence, reduces the pressure on the health care system. In the fight against the virus, social distancing had already been known to be an efficient method of limiting disease spread. People are being

P. Baskaran (✉) · R. Rengarajan · S. Naveen Kumar · V. Vijay
Easwari Engineering College, Chennai 600089, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_20

told to limit their social interaction with one another in order to reduce the likelihood of the virus spreading through bodily or close contact.

According to the World Health Organization, virus is transferred via drops and feces during nearer unsecured interaction with an infector and an infected person. As a result, staying away from people and touching infected patients' belongings are often used to avoid infection transmission. AI/deep learning has previously demonstrated positive results on a wide range of lifestyle issues. This developed scheme will include a detailed description of how we will utilize the available technologies to monitor social distance in public areas and work spaces.

To make sure that the social distance and all other protocols are being followed, it is necessary to monitor the people if they are following or not. This tool can be used in such scenarios to avoid human monitoring (which is not advised in this pandemic situation). This camera live video or the already taken video can be fed to the tool, the tool analyzes the video and monitors the people and check if social distancing is being followed or not. This tool can be integrated to any sort of monitoring or security system.

The pandemic is not over yet, and many countries (like India, South Korea, Australia) are experiencing multiple waves of virus spreading. This is highly due to the lethargicness of people not following the protocols and country's economy. Although the vaccines can reduce the probability of getting affected, the availability is rare since there is a huge demand and the production of the vaccine is slower. To minimize these risks, it is recommended that people stop any contact, such as holding hands, and keep a minimum distance of 1 m between themselves.

It is risky to get out and shop for the necessary items during this pandemic as we will expose ourselves to the public which will increase the chance of getting affected by the virus. However, it is advised to follow some strict protocols and to implement monitoring tools to check. So, once you exit during the pandemic, follow social distancing together from the only essential safety rules at shopping malls. You will notice that even the security at the malls would ask you to maintain distance, even if you are going with someone you know or are close to. Points that summarize the important components of this approach are: DL has gained prominence in object recognition that has been used for image processing; a tool that has been developed to compute the distance between the people and alert accordingly; live video analysis and categorization of objects; importance of maintaining the distance.

2 Literature Survey

Accurate localization may be a critical requirement for synchronized localization and mapping (SLAM). SLAM is a mapping technique which is used to classify the objects in the immovable or static environment [1]. It is not possible to process the movable or dynamic environment using this technique. Although there are a variety of SLAM [2] frameworks, they also perform poorly due to low localization accuracy.

Here, a new SLAM named RGB—D SLAM [3] which stands for red, green, blue—depth simultaneous localization and mapping is created to improve the localization accuracy.

Shirmohammadi and Ferrero [4], we continue to have access to modern technologies as both hardware and software components in image sensor and processing platforms progress at a fast rate. Scientists and professionals are increasingly being used as clichéd tools to continue living and supervise physical phenomena, among other things. This article examines this growing trend, as well as how sensors and sight are being used for instrumentation and measurement. The business process unit receives the visual sensor's image and carries out the basic ops to attain the given metrics. This unit can be enacted in any software or hardware; for example, it can be coded into a clichéd microchip system, such as the processing system of a camera sensor, or it is frequently put into practice in devoted hardware, such as field programmable gate array (FPGA) or application-specific microcircuitry (ASIC). Furthermore, it would not be in the format required by subsequent operations.

Sun et al. [3], SLAM is a mapping technique which is used to classify the objects in the immovable or static environment. It is not possible to process the movable or dynamic environment using this technique. Although there are a variety of SLAM frameworks, they also perform poorly due to low localization accuracy. SLAM is an open-source mapping technique. The suggested technique functioned as which was before stage, filtering data involved in moving objects. We put our method to the test using a publicly available RGB-D dataset [5]. The simulation results reveal that our policy was capable of significantly improving RGB-D SLAM efficiency. This work is an expanded adaptation that attempts to work out the RGB-D classification problems in dynamic environments.

Mur-Artal et al. [1], ORB-SLAM is an app monocular SLAM process to manage in real time in undersized and outsized indoor and outdoor. The arrangement is resistant to stern movement clutter, has a broad benchmark recursion and delocalization capability, and embraces complete usual initialization. We created a one-of-a-kind system which uses equivalent features for all SLAM tasks. The best approach for selecting the reconstruction's points suggests that the current results in excellent reliability and creates a compressed and findable map which only develops if the episode changes, allowing for lifelong operation. ORB-SLAM outperforms other modern monocular SLAM advances in terms of performance. Our previous work is supported by the loop closing and delocalization methods presented at this point. The system's preliminary version was unveiled. This paper introduced the preprocessing of live video, monitoring, and social distance analysis together. It also clearly depicts collection of how the social distance is measured in real time through experimental scheme. Also, the best part of this anticipated model is that the attributes used for mapping and tracking are used for position identification in order to perform frame rate re-localization and loop avoidance. This improves the efficiency of our scheme and eliminates the necessity to interpose the depth of popular characteristics from nearby SLAM features, as in prior efforts. Here, from the most of the surveyed works, it is evident that it does not concentrate on moving object detection [6] and alerting in a dynamic environment. So, for the urge of handling and controlling over the

pandemic situation, it is essential to propose an automatic social distance regulating and alerting system [7] for helping the mankind.

3 Proposed System

The figure shows the functional architecture of the system. It tells about the complete flow of the system. Right from the preprocessing to sending, an alarm is designed in the functional architecture itself. There are three modules in the system, and it is well explained and much needed for the system to work flawlessly. It is a modular design so that the module can be re-structured or re-used for similar systems like this. In the first module, that is, the preprocessing module, the frame conversion and background subtraction take place [8], and further, the human objects are detected in the initial module itself. In the second phase, the detected human objects [9] are tracked and boxes are bound around them. In the third and final phase, the distance between the boxes is calculated, and if the distance is lower than the threshold, that is, high risk [10], the alarm is raised and a number of high risk people is calculated and displayed on the screen. If the distance calculated is higher than the threshold, that is, low risk, nothing has to be done. The proposed system as shown in the below Fig. 1 is a modular design so that the components of the system can be separated and joined if needed. Modular design always proposes a positive impact due to agility in developing and maintaining the system in the long run [11].

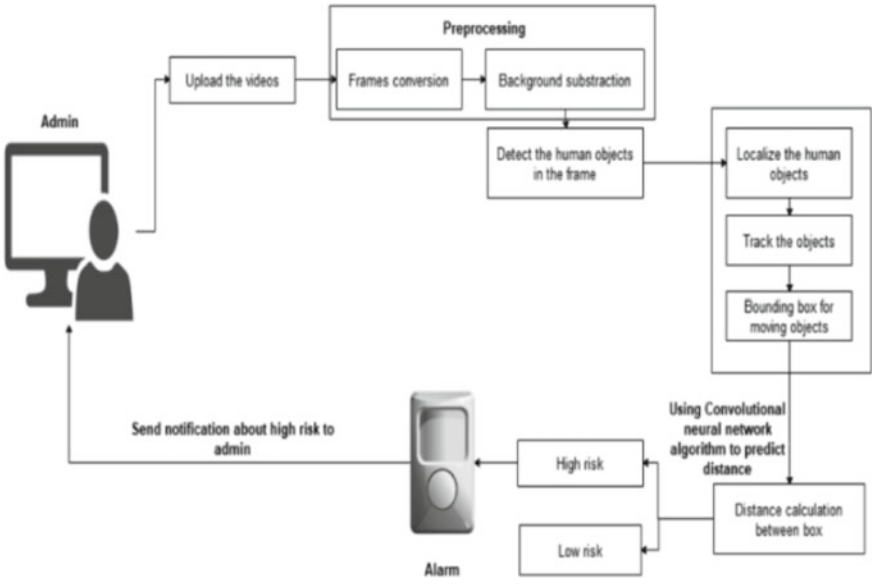
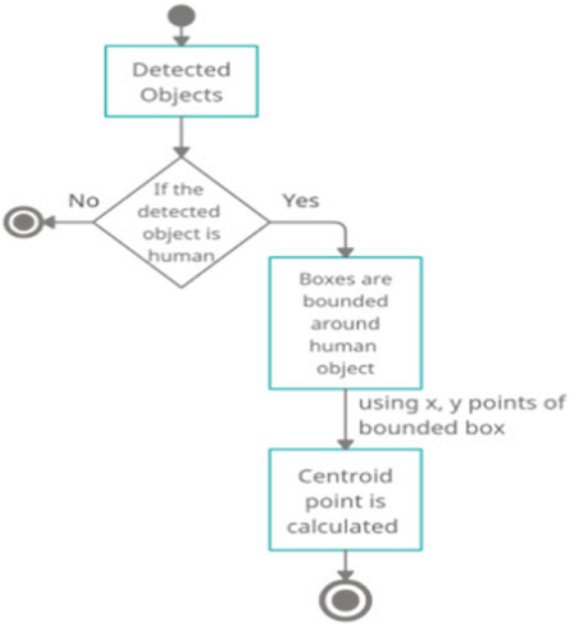


Fig. 1 Proposed architecture

Fig. 2 Preprocessing



The system consists of the following modules.

3.1 Preprocessing

Based on the flow in Fig. 2, the camera is fixed at an angle to capture the video which is feeded to the system. For pre-captured video: The video is directly feeded to the system. Captured video frames are input to the system and convert the videos into frames at 0.5 s. Moving object is detected using you only look once (YOLO) algorithm.

3.2 Tracking and Monitoring

The following steps as shown in Fig. 3 will happen for drawing a box over each individual. To reduce the risk of over fitting, we use deep learning and various norm heuristic—based to clean up the output bounding boxes. Bounding boxes are constantly tracked and monitored [12].

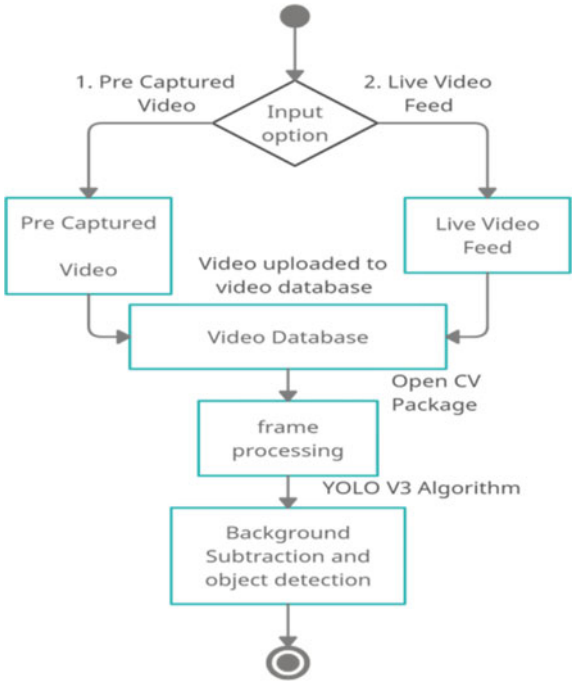


Fig. 3 Tracking and monitoring

3.3 Distance Calculation

Each and every frame is then converted into bird's eye view (Fig. 4), i.e., two-dimensional top view using Python OpenCV2 library. Then, distance is calculated using the centroid point of every human individual. If the calculated distance between every individual is higher than the threshold (6 feet), NO alarm is raised. If the distance between the individuals is lower than the threshold, then they are considered as a high risk individual and alarm is raised.

4 Results

This implementation is developed with the help of Python with deep learning approach using OpenCV, and below pictures in Figs. 5 and 6 show the working screenshot of the proposed system.

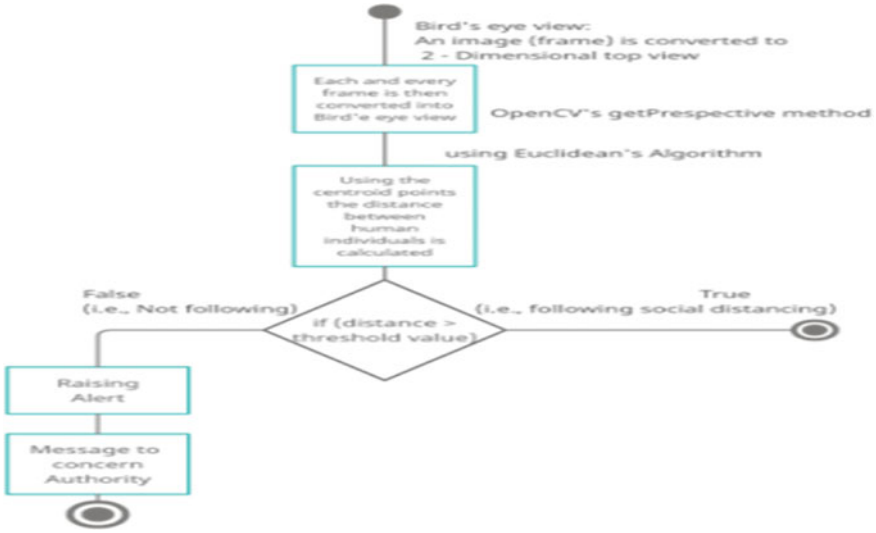


Fig. 4 Distance calculation

Fig. 5 Sample for consideration and preprocessing stage

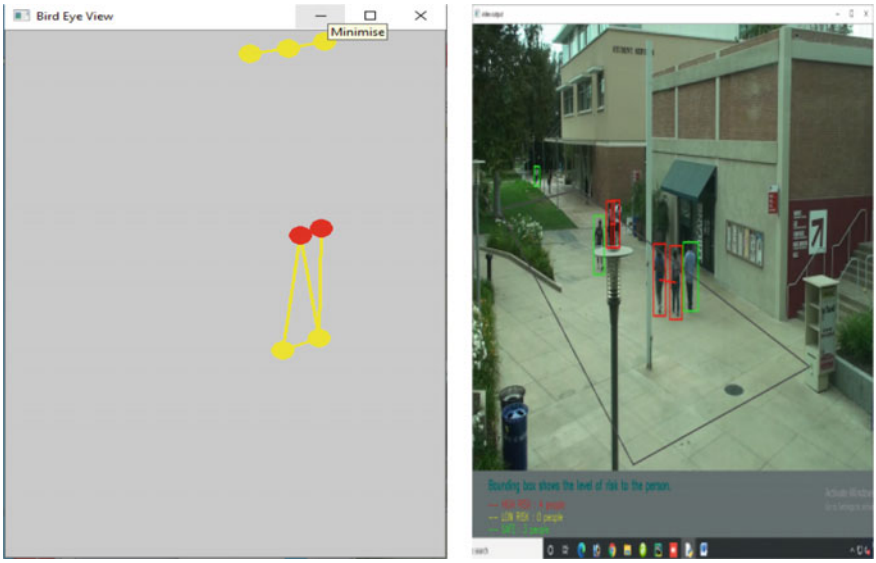


Fig. 6 Distance analysis phase and risk identification

5 Conclusion and Future Enhancement

The proposed system can be used to reduce the risk in spreading the virus by monitoring with a computer-based system. As there are already surveillance systems available in every public place, this tool can easily be integrated to the existing system thus reducing the extra cost. Also, this reduces work of manpower by automatically doing the task on its own. These types of tools can be implemented in all types of places with basic hardware and software. As this system or tool can have every corner covered up for the searching and monitoring, the accuracy is a bit higher than some of the other tools. This tool uses some powerful background subtraction method to differentiate human objects from similar human objects like chimpanzees. The result from the experimental test shows that the tool can produce better results when used in a real-world testing environment. Further, in the future, a variety of datasets ranging from mid-large to large size can be used for training to improve the accuracy. And, many modules like sanitizer switch (to check if the people have used sanitizer before entering), face mask detection can be implemented in the same tool as it is a modular tool.

References

1. Mur-Artal R, Montiel JMM, Tardos JD (2015) ORB-SLAM: a versatile and accurate monocular SLAM system. IEEE Trans Rob 31(5):1147–1163

2. Xie W, Liu PX, Zheng M (2021) Moving object segmentation and detection for robust RGBD-SLAM in dynamic environments. IEEE Trans Instrum Meas 70:1–8
3. Sun Y, Liu M, Max Q-H, Meng (2018) Motion removal for reliable RGB-D SLAM in dynamic environments. Rob Autonomous Syst 108: 115–128
4. Shirmohammadi S, Ferrero A (2014) Camera as the instrument: the rising trend of vision based measurement. IEEE Instrum Meas Mag 17(3):41–47
5. Wang Y, Huang S (2014) Towards dense moving object segmentation based robust dense RGB-D SLAM in dynamic scenarios. In: 2014 13th international conference on control automation robotics and vision (ICARCV). IEEE
6. Ochs P, Malik J, Brox T (2013) Segmentation of moving objects by long term video analysis. IEEE Trans Pattern Anal Mach Intell 36(6):1187–1200
7. Baskaran P, Aasim Ahmed SR, Yusuf A, Bordoloi D (2020) Smart recognition and apprise system With IoT. Int J Cont Autom 13(02):604–610
8. Cui L, Ma C (2019) SOF-SLAM: a semantic visual SLAM for dynamic environments. IEEE Access 7:166528–166539
9. Rohith S, Baskaran P, LMNKVNRR (2021) Inshore and offshore ship detection and categorization system using mask R-CNN. Des Eng 529–541
10. Xu B et al. (2019) Mid-fusion: Octree-based object-level multi-instance dynamic slam. In: 2019 International Conference on Robotics and Automation (ICRA). IEEE
11. Baskaran P, Karuppasamy K (2022) Hybrid teaching learning approach for improving network lifetime in wireless sensor networks. CMC-Comput Mater Cont 70(1):1975–1992
12. Baskaran P, Anurag EG, Kaviyarasu S, Eniyavan AV, Karthikeyan M (2017)Time slot based data sharing with time slot based password. In: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp 1–3. https://doi.org/10.1109/ICIIECS.2017.8276146

Critical Data Capturing of Dual Mobile Data Collectors in WSN

Sunita S. Patil and T. Senthil Kumaran

Abstract Data gathering infrastructure is greatly necessitated once significant information from big data systems are discovered. Wireless Sensor Networks (WSNs) are regarded as powerful data sources, i.e., huge data generation is done by several nodes in large-scale networks amid various feasible data sources. Conversely, WSNs suffers from severe deficits pertaining to data reliability and communication due to nodes restricted abilities. Furthermore, in the deployment of enormous nodes during a dense manner, the quantity of sensed data is not considered. Sensor nodes in network may need to communicate the gathered info by stipulated time, but it is delayed due to gather or communication, hence, degrade the performance. This is routine problem, to encounter the improper energy efficiency, the mobile data collectors are performed. The proposed used two mobile data collectors, one for essential info collector and second one of normal info collector. Hence, the couple of nodes will not schedule in data collection from collector trajectory as their stipulated time which is decreased than the entire travel time of the data aggregator, therefore such cluster heads are isolated of the trail, this should cause a greater number of cluster heads to remain in path in future iterations, also it must have very little effect on the tour time. It is thereby validated that suggested methodologies attain the necessitated reliability along with network lifetime maximization by means of simulation output when compared with prevailing methodologies.

Keywords Data collector · Sink · Critical data collector

1 Introduction

Recently, Wireless Sensor Network plays a serious role within the information gathering paradigm [1]. There is no pre-configured infrastructure for sensor nodes deployed in large-scale sensing field. The nearby nodes are identified for organizing themselves into a network once the environment is monitored. On the

S. S. Patil (✉) · T. Senthil Kumaran
ACS College of Engineering, VTU, Bangalore, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_21

other hand, network lifetime is especially determined by an information gathering plan [2]. A standard feature is shared generally by them, albeit sensor applications could additionally be relatively unique: The total of their information packets is completed at some information sink. Detecting primarily depends on rate regardless of being hooked in to sensor area/network topology, it is reasonably constant. Additionally, information gathering procedure has a prominent part in characterizing the organization lifetime [2]. Thus, fizzling of sensors may cause difficulty in achieving the collector and subsequently the network gets disengaged, even the vast majority of the hubs may calm endure for an extended period. Within the case of an enormous scope information driven sensor network, one static information sink is not efficient for information gathering from all sensors. Be that as it may, in couple of utilization, sensor deployment is completed for observing separate regions. In each territory, sensor thick organizations are associated, yet sensor sending having a place with divergent areas could likewise be disconnected. As like completely associated networks, certain sensors are unable in information sending to information sink through remote connections. A portable information gatherer is easily fitting to those applications. A versatile information collector is a portable "information carrier" which helps in moving by means of every local area other than connecting all disconnected sub-networks together. The portable information authority moving path is virtual joins in the midst of isolates sub-organizations. Different chains of command to networks are introduced in before work [3, 4] due to flat geography limitations. In those organizations, sensor hubs organization is completed into bunches which comprise the organization lower layer. At the upper layer, detecting information from sensors are gathered through bunch heads, and information forwarding is completed to the surface data sink. The equivalent capability and energy at the start are controlled by all nodes during a homogeneous network during which few hubs are picked to proceed as bunch heads [5]. Yet, nearly higher energy is devoured by group heads than other sensor nodes. So as to moderate the bunch head fizzling problem before other hubs, sensor nodes are often utilized as group heads during a rotational way [6]. In such a network, thanks to the likelihood of each sensor hub being a group head, each sensor hub is adequately solid for dealing with approaching and active traffic and reserve detecting data, which can increase the whole sensor organization's general expense. Also, group head selection during a dynamic way may cause high overhead due to recurrent data trade in the midst of sensor hubs. The critical weakness of homogeneous networks is often enhanced by presenting a little number of asset rich hubs. As opposed to homogeneous organizations, a heterogeneous sensor network envelops not many asset rich nodes alongside an enormous number of asset restricted fundamental sensor nodes. There is only limited correspondence capability and mostly accentuation on climate detecting, though asset rich hubs equipment is completed with all the more impressive handsets and batteries. There will be two sorts of mobile authorities, one for basic information collection and accordingly the other for non—basic information collection. It would be conceivable that a couple of nodes will not slot in the versatile gatherer way as their cutoff time is lesser than the entire tour-season of the portable gatherer, thus such bunch heads are isolated from the trail, this could cause more number of group heads to stay within

the path in future cycles, likewise it must have little or no effect on the visit time [7, 8].

2 Literature Survey

There are various existing energy protocols that are analyzed are often explained below.

2.1 LEACH Protocol

Low Energy Adaptive Clustering Hierarchy this convention was proposed by Hein Zelman. During this convention, the sensor hubs during an organization are assembled under bunches. Each group is allocated a bunch head that assembles the data from its bunch prior to sending it to the base station. It is been suggested that lone 5% of the whole hubs should become group sets out toward more prominent effectiveness [8, 9].

2.2 Mobile Data Collectors in WSN

The portable information authorities visit each bunch head and accumulate information from them and visits back to put together station. This manner the overhead with respect to the group sets out toward sending information to the bottom station is removed leading to these hubs to survive for extended. The exploration was distributed while improving the productivity as expected for information assortment by coordinating different reception apparatuses onto the portable authority. The locations are precluded from the list leading to the rundown of surveying focuses where versatile authority would stop to collect the information [1, 2].

Figure 1 shows a representation of the LBC-DDU (Load adjusted grouping and double information transferring) system.

Layer—1: This layer depicts the trail taken by the portable gatherers among the varied polling focuses accessible.

Layer—2: This layer portrays the group heads, at the most two bunch heads are often appointed by each limits framed by the groups.

Layer—3: This portrays the actual layer where all the hubs are sent, it likewise portrays in the very paper the creators have broken down the assortment of information with time requirement (i.e., basic or postponement sensitive).

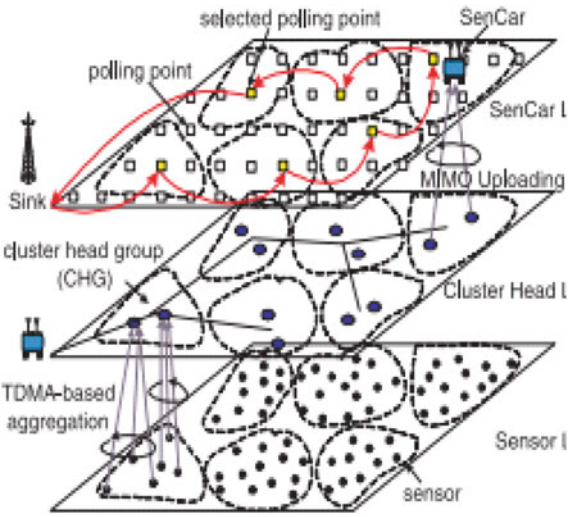


Fig. 1 Illustration of LBC-DDU framework

2.3 Tour Planning for Mobile Data Collector

This algorithm is aimed to identify the tour through single hop polling points, the neighbor set of, neighbor set is defined as a primed device which transfer aggregation directly to the mobile aggregator deprived of any relay, denotes set of uncovered neighbors as defined as the cost of reaching the neighbor set from any previously covered neighbor set [8].

2.4 Dynamic Deadline in WSN W*as also Investigated* U*pon*

Table 1 summarizes briefly the literature survey for the work and subsequently by examining the current methodologies; the exploration work commitment contains the resulting factors. The helpful information assortment issue by versatile authorities is planned as coordinating issue in the midst of portable gatherers and significant level bunch heads. Their advantage positioning networks are developed dependent on portable authorities' inclinations and recommended a coordinated coordinating calculation [9].

3 Proposed Methodology

The proposed system employs two mobile collectors. The mobile aggregator for collecting the information of cluster heads be classified as censorious and the other

Table 1 Summary of literature survey

Title	Objective	Advantage	Drawbacks
Time orient LEAD-based polling point selection algorithm for efficient data aggregation in wireless sensor networks [10]	This paper presents a time orient location energy availability data rate (LEAD)-based scheduling algorithm	TLEAD outperforms MSPR by 46% regarding delay, by 16% for delivery ratio, by 50% for packet drop and by 7% for residual energy	Improvements regarding critical data and security model
An optimal data gathering method for mobile sinks in WSNs [9]	K—means algorithm is used for clustering	The network lifetime is increased to 20% when sensor nodes are divided into from 4 to 15 clusters	Does not mention about the critical data and the path for data collection is not optimal
Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks [1]	Uses MU-MIMO data collector and collects the data from the cluster heads formed	In theory, with MU-MIMO the data collection can be reduced to almost half	Improvements is polling point selection was suggested
Tour planning for mobile data gathering mechanisms in wireless sensor [8]	An M-collector gathers data form static sensors at the polling points, it also proposes deployment of multiple M-collectors	Considers the obstacles in data transmission at polling points and checks the same during initialization	Overlapping of sensor nodes is ignored and delay sensitive WSN topology is not considered

mobile aggregator is engaged to gather from the cluster heads be categorized as noncensorious.

3.1 Network Model and Assumptions

In Fig. 2 uncovers the organization model considered in this exploration. The succeeding are the speculations in this exploration: An occasional information gathering application is considered in which detected occasion bundles are sent by sensor hubs intermittently to their closest DCs through multi-bounce correspondence. Each DC contains sensor hubs group answering to it. All sensor nodes are considered as static and positioned consistently in the specified terrain. All DCs are said to have development capacities and are more intense than sensor hubs relating to battery energy, correspondence, calculation, and memory assets. Afterward, this supposition that is loosened up that all the DCs are portable, i.e., few are static. Similarly, it is accepted that no obstruction obstacles are available for DCs development. The occasion parcels are collected by DC got from the unmistakable sensor hubs and these

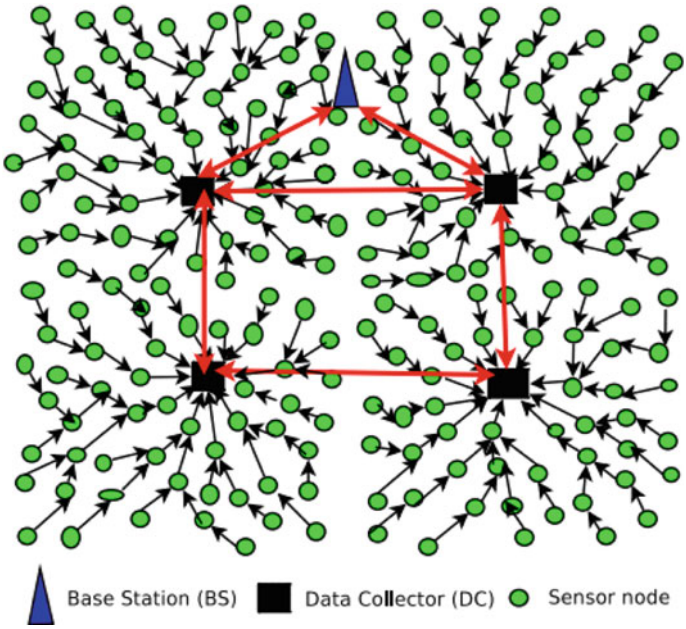


Fig. 2 An example of network model

total occasion bundles are accounted for to BS through multi-jump correspondence over different DCs. The DCs should be static nearly and travel from one situation to the next simply after a choice.

3.2 Proposed Data Collection

The grouping of the organization into basic and non-basic should be possible dependent on the cutoff time and the precession required, the method is talked about in detail in the outcomes and conversation area underneath. In the proposed framework the bunch head gathers cutoff time data from all its part hubs other than registers the normal of the equivalent. Consequently, registered normal is sent to base station which will run calculation just as bunches division into two classes is done: basic and non-basic. Any of the bunch heads whose cutoff time is lesser than the absolute visit season of the versatile hub is taken out from the way and will currently be classified as non-basic despites of its being basic. Just such bunch heads should be eliminated which would cause the recovery of more number of group heads in the visit during future emphases, likewise the absolute visit time (the visit through the portable gatherer is acquired by the calculation Traveling Sales Person—branch and bound) of the versatile authority can be diminished by just a little which can help in accounting for more group heads to stay in way during future cycle (more modest

the decrease in visit time—more is number of bunch heads that can stay in visit). The methodology is proceeded as long as there are bunch heads that do not consent to the visit. Therefore toward the end just such group heads stay in the visit whose request to satisfy the cutoff time is conceivable. The group heads eliminated from this rundown are appointed to another portable gatherer who gathers from non-basic bunch heads.

The above method is recreated without fail while taking out the group heads individually until just 4 bunch heads are left in the rundown, in this way gotten consequence which is dissected, hence a choice about the order is ready-made dependent connected to the prepossession and the concealment needed.

Algorithm:

```
While (true){
    # identify the clusters and the respective cluster heads as CH
    CH = identify_clusters(Network_Topology);
    #identify delay sensitive cluster heads (DSCH)
    # assign to MC-1 the poling points "CH-DSCH," that is non-prominent CHs
    MC-1 = {CH-DSCH};
    Flag = true;
    While(flag) {
    # find such that the deadline (&& deadline is least of all deadlines that are identified
so far.
    List_DSCH = all where (deadline(holds;
    If(list_DSCH = = null) {
    Flag = false;
    Continue;
    }
    To_remove_cluster_i = one_with_least Random_deadline(list_DSCH);
    #the unqualified multi-hops its data directly to BS
    MC-1 = {to_remove_cluster_i} or multihop_to_BS(to_remove_cluster_i);
    DSCH = DSCH-to_remove_cluster_i;
    #compute the new TOUR_TIME
    TOUR_TIME = new_TOUR_TIME(Dsch);
    }
    #thus computed list is assigned to MC-2
    MC-2 = {list_DSCH};
    # commence and complete the tour
    }.
```

4 Results and Discussion

Following outcomes are determined on running the above calculation for around 15 bunch heads. The exploratory outcomes are reproduced utilizing Ns-2 SIMULATOR. Three distinct assortments of the outcomes are examined. As referenced in the framework outline segment where it is referenced that any of the group heads which do not consent to the visit time is taken out from the way, the expulsion of the hub can be founded on after measures: (1) irregular bunch head (2) cluster head with greatest cutoff time first, and (3) bunch head with least cutoff time first. From the acquired outcomes portrayed in Figs. 3, 4, and 5 it is hence obvious that the evacuation of group heads dependent on maximum cutoff time first when not agreeable yields better outcomes by arriving at 100% a lot speedier than the previous two expulsion strategies. It is conceivable that other two techniques can be useful during different situations (i.e., least cutoff time first and irregular expulsion strategies can likewise yield better outcomes in different situations). To characterize the groups as either basic or non-basic is done dependent on the exactness and the cover required, assume for the accuracy of 100% one may get a front of 7 bunch heads under basic versatile authority. The above chart Fig. 6 portrays the effectiveness of the current framework, where the second any infringement of the cutoff time happens the versatile authority stops the visit and returns back to base station for basic information conveyance, it at that point returns to similar position and continues the visit there onwards, the disadvantage of the framework is that it neglects to convey 100% productivity for quite a few supporting bunch heads, it delivers a proficiency of at max 83% just when number of group heads is 12, not normal for which proposed

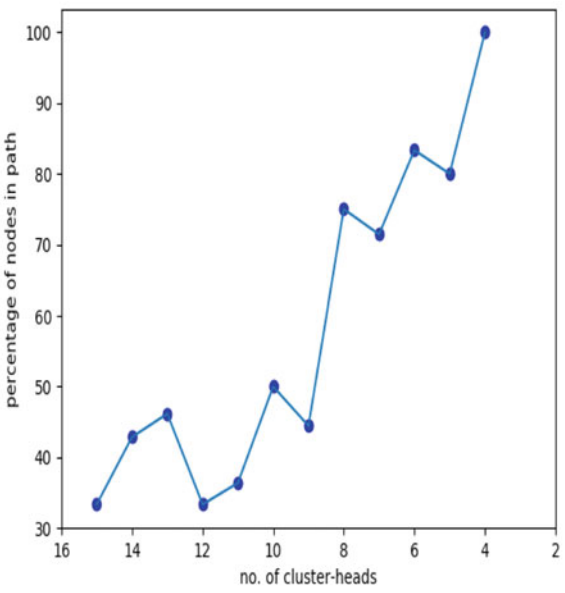


Fig. 3 Random cluster heads which are incompatible with the tour time are being removed

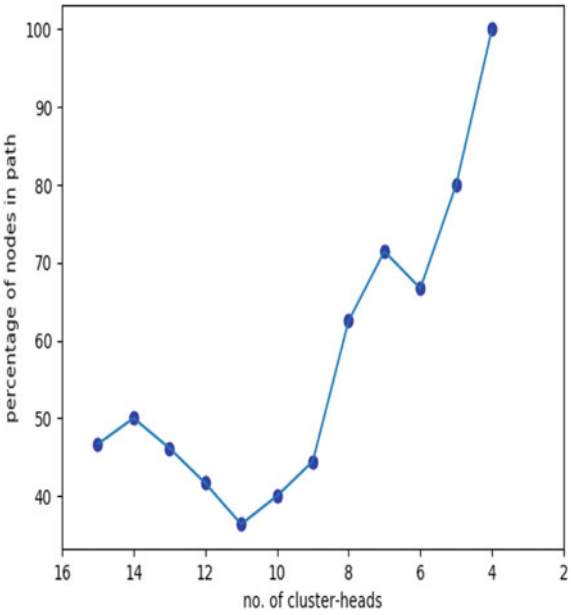


Fig. 4 Cluster heads with least deadline first being removed

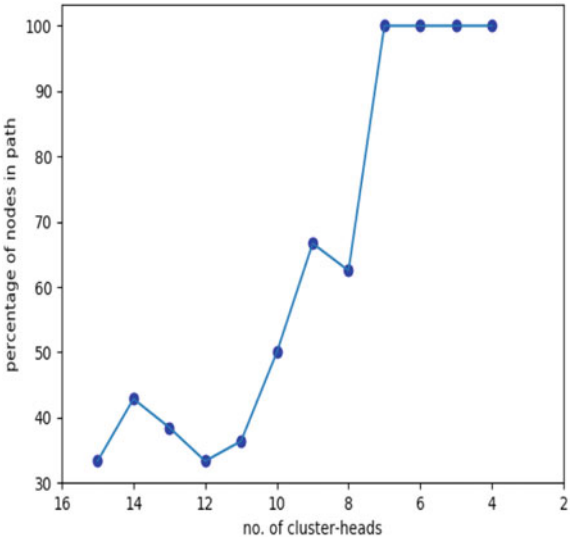


Fig. 5 Cluster heads with maximum deadline first being removed

framework yields a 100% cover in any event once in any situation as obvious from the above demonstrated diagrams. Figure 7 shows the diagram portrays the time utilization of both existing and proposed framework, obviously the current framework devours somewhat lesser time in contrast with proposed framework, the red line demonstrates the time devoured by proposed calculation which is 1.25 for 14

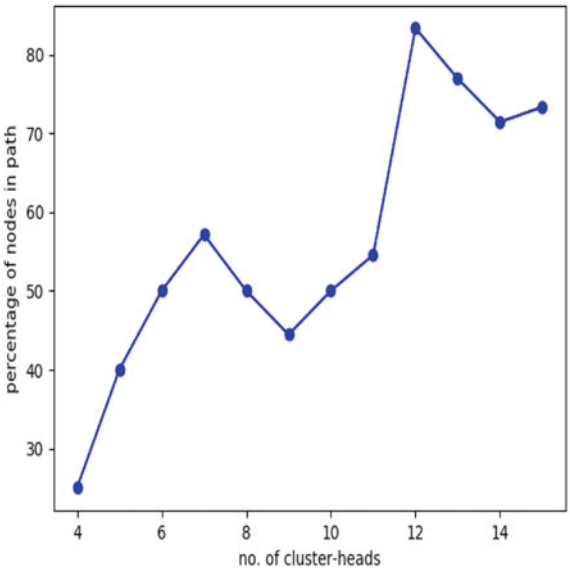


Fig. 6 Outcome of earliest deadline first algorithm

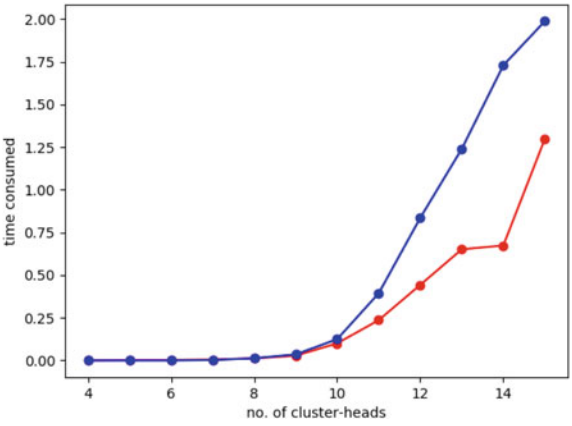


Fig. 7 Time consumption versus number of cluster heads

no. of group heads and the blue is for existing framework devours additional season of 2.00 for 14 no. of group heads.

It is interested to see that current framework accomplishes better once number of bunch heads are more noteworthy, whereas the proposed framework accomplishes better similarly number of group heads are tallied downwards, it is to see that the proposed framework has beated existing framework in lion's share of the cases as we lessen the tally of number of bunch heads.

5 Conclusion

There had consistently been a surprising advancement in remote sensor network space about the boundaries of energy and group development however barely a couple of papers are devoted to a similarly significant boundary which is basic information. The creators of article is likewise quick for seeing whether a couple of the bunch heads by exceptionally basic forward node information in multi-jump to the base station straightforwardly non-trusting that any versatile authority will meeting it for information assortment. For example, the time devoured by proposed calculation which is 1.25 for 14 no. of bunch heads and the blue is for existing framework devours additional season of 2.00 for 14 no. of bunch heads. The creators additionally recommend that in the event that insight could be fused into the organization, at that point the basic information regardless of whether dynamic can be better taken care of despite this paper recommending an ideal calculation to deal with basic information the creators urge the endeavors to expand a similar when the criticalness is dynamic in future.

References

1. Zhao M, Yang Y, Wang C (2015) Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks. IEEE Trans Mob Comput 14(4)
2. Mathapati M, Kumaran TS, Muruganandham A et al (2020) Secure routing scheme with multi-dimensional trust evaluation for wireless sensor network. J Ambient Intell Human Comput
3. Arora VK, Sharma V, Sachdeva M (2019) ACO optimized self-organized tree-based energy balance algorithm for wireless sensor network. J Ambient Intell Humaniz Comput 10(12):4963–4975
4. Heinzelman WR, Chandrakasan A, Balakrishnan H (2000) Energy efficient communication protocols for wireless microsensor networks. Proc HICSS, Maui, Hawai
5. Younis O, Fahmy S (2004) Distributed clustering in ad-hoc sensor networks: a hybrid, energy-efficient approach. IEEE INFOCOM 2004, Hong Kong
6. Amis A et al (2000) Max-min D-cluster formation in wireless ad hoc networks. IEEE INFOCOM 2000, Tel-Aviv, Israel
7. Zhang Z, Ma M, Yang Y (2008) Energy efficient multi-hop polling in clusters of two-layered heterogeneous sensor networks. IEEE Trans Comput 53(2):231–245
8. Ma M, Yang Y, Zhao M (2013) Tour planning for mobile data-gathering mechanisms in wireless sensor networks. IEEE Trans Veh Technol 62(4):1472–1483
9. Al Islam AA, Hyder CS, Kabir H, Naznin M (2010) Finding the optimal percentage of cluster heads from a new and complete mathematical model on leach. Wirel Sens Netw 2(02):1–129
10. Yuvaraj R, Chandrasekar A, Jothi S (2019) Time orient LEAD based polling point selection algorithm for efficient data aggregation in wireless sensor networks. Cluster Comput 22:3339–3346

Improved Congestion Control in Wireless Sensor Networks Using Meta-Heuristic Approach

Kavita K. Patil, T. Senthil Kumaran, and A. Y. Prasad

Abstract The wireless sensor network facilitates a wide range of sensor nodes and allows various real-time applications. The problem of routing is a very important in wireless sensor networks that need to be taken into consideration to maximize network life. As the network is grows, routing become more complex. A new congestion control system is proposed at optimised rate in order to provide energy-efficient transmissions. This involves the rate-based cluster routing congestion management algorithm to reduce energy use across the entire network. The rate control method reduces end-to-end time for the wider simulation cycle, to increase the lifetime of the network. Initially, the nodes are clustered with the K-means, with optimization of Bee Swarm for high packet delivery ratio and Particle Swarm Optimization-based routing for packets with maximum throughput. The experimental results shows that the proposed method achieves higher rate of throughput and reduced control overhead than other methods.

Keywords K-means clustering · Particle swarm optimisation · Bees swarm optimisation · Congestion control

1 Introduction

The wireless sensor network consists of a variety of powerful and meaningful sensors, which capture and relay information from your neighbour. Sensor networks, for example bandwidth, battery capacity, available memory and data speeds, are reserved for various resource reservations. Different sensing and monitoring systems such as defence, industry output and traffic management can be used by the sensor function

K. K. Patil
ACS College of Engineering, Bangalore, Karnataka, India

T. Senthil Kumaran
ACS College of Engineering, Bangalore, Karnataka, India

A. Y. Prasad (✉)
SJB Institute of Technology, Bangalore, Karnataka, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_22

[1]. Dependent on the request, full, unexpected and related information signals will yield effects that can be transmitted to a minimum number of sinks. The effect is packet crash or overflowed buffer in the channel [2]. This causes congestion. Some methods of WSN management are based on the following conditions [3–5]: quantity, feedback sort, deployment power, efficiency and fairness. The following methods can be used to control the congestion.

In addition, one of the key limitations for WSN is its energy consumption. Several methods are being applied to increase energy consumption to eliminate data errors when transmitting network life with improvement [3]. A congestion-prevention approach known as hierarchical congestion management uses lower adjacent hierarchy nodes to transfer data. Next, the congestion management queue length is calculated [4]. It did not dwell on energy consumption at lower hierarchical nodes as the downside of this strategy. Congested nodes based on Relative Success Rate (RSR) values [5–8] are avoided by choosing the next hop node.

2 Related Work

6LoWPAN structures have been seen here a basic congest management mechanism called an IEEE 802.15.4 game theory-based congestion control application. It supports application scenarios dependent on goals for the node and application. This approach determines the suitable propagation rate for all leaf nodes. In additive-increase–decrease congestion management, the delivery rate of cognitive radio sensors is modelled with the distribution by a semi-Markov chain of a cognitive radio node by assigning functions of the transmission of sensors and input speeds. The key impacts of the device include consumer versatility in the allocation of transmission speeds [8].

The paper focuses on developing a high-life energy-efficient network system. In this article, we establish a new mechanism to avoid the congestion of the adaptive rate and its consequent problems in the WSN, which are still present in prior methods. It means that the end-to-end delay and the loss rate between the sensors and the sink node are reduced.

Zhuang et al. [9] has proposed different techniques to reduce to a network congestion in the minimum path can be different and different classifications used for MANET routing algorithms: proactive routing against request or rispetto single routing multipoint routing in proactive protocols, paths are between nodes due in early, though transmission and transmission is not required. Appropriate approach for non-large dimensions nets, because many tracks not to be used yet, they must keep up and regularly update an overwhelming suv in head treatment and communication. The more efficient approach to what discovered travel track only when needed for transmission and transmission is released when it is no longer happening. Table 1 show the Comparison of Congestion Control Techniques based on Priority.

Table 1 Comparison of congestion control techniques based on priority

Protocols	Congestion			Methods	Performance metrics
	Detection	Notification	Avoidance		
PCCP [11]	Intelligent Congestion Detection (ICD)	Indirect	Packet rate readjustment	Introduced node priority	Regularized throughput, buffersize and weighted fairness
Quality of Service Adaptive Congestion Control [12]	Congestion scale	Indirect	Packet rate readjustment, with alternate route	Adaptive cross layer approach to QoS	Average buffer occupancy, Average leftover energy
PHTCCP [13]	Packet service ratio	Indirect	Packet rate readjustment	Inter and intra-queue priorities	Regularized throughput, delay, energy efficiency
Priority based rate control for service differentiation [14]	Congestion index	Indirect	Packet rate readjustment	Differentiate between real-time and non-real-time traffic and serve them accordingly	regularized throughput, loss probability, delay
ECODA [15]	Buffer, Weighted buffer difference	Indirect	Packet rate readjustment, multiple route	Priority is set dynamically	Throughput, delay and queue occupancy
Priority based congestion control for multipath [16]	Packet service ratio	Indirect	Packet rate readjustment, multiple route	Considered in multipath circumstances	Regularized throughput using fairness
PNBCC [17]	Ratio of current buffer length and current queue length	Indirect	Packet rate readjustment	Identify priority node by length of queue	Throughput, packet loss, channel occupancy time, average delay time

3 Proposed Methodology and Description

The new rate regulation based on greedy clustering and colony routing is the subject of the proposed process. It has used many sources and a drain. Network nodes are initially grouped using the clustering algorithm of K-means. Each node dispenses its distance from another node and energy over the whole community of nodes. For the efficiency of the device, the sum of residual energy and retransmission is important.

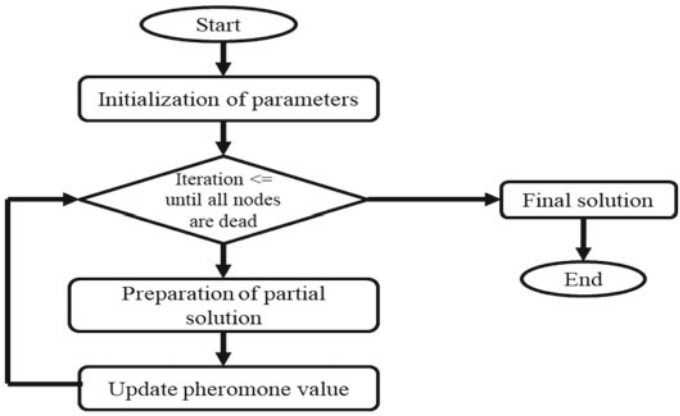


Fig. 1 ACO routing

Tracking packet flow from several outlets to sink node is the basis for network congestion. The queue length is an important parameter here for tracking congestion. The queue length size is spread throughout the system. During this process. Data packet networks are used to path the window size and queuing size (using ACO). The buffer size is checked as the packet sends from the source to another node: reaching the buffer size means transmitting it as zero. The packet rate is reduced by the BSO if the accreditation is zero. Each node sends data to another cluster node. The routing table must be updated by the continuous queue size of other knots. The key principle here is to maximise the rate in order to collect full data at the recipient (Figs. 1 and 2).

Clustering phase: Transmission consists of transmission of packets between nodes and heads of clusters. We add the clustering nodes K-means algorithm. Initially, this algorithm finds the middle and measures the distance between each node. A cluster is finally formed from this.

Packet rate reduction: Indicator of congestion is submitted by means of an acknowledgement. In the case of a positive affirmation, the packet rate must not be limited to a minimum. In a second example, the BSO algorithm reduces packet speed. The key goal is to minimise the rate of the packet.

Enhancing the throughput: The PSO routing improves performance by updating pheromone values by the objective feature in any routing direction. An arbitrary direction to update smaller distances is chosen on the basis of a pheromone to minimise processing time. The lowest direction for transmitting information must be sought by PSO. It was simulated from one point to another using knowledge from the adjacent ants and a number of the ants. An ant follows its precedent ant-pheromone set in the algorithm and acceptable path predictions are based on the pheromone values.

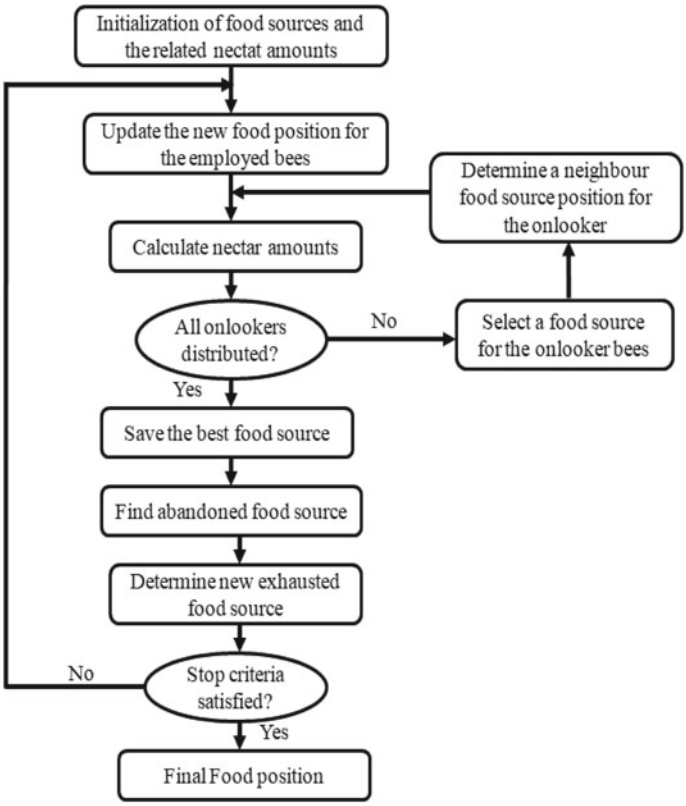


Fig. 2 BSO optimisation

4 Results and Discussion

The proposed BSO–PSO is compared with other existing methods that includes hybrid multi-objective optimization (HMOO) algorithm [10], accuracy-aware congestion control (AACC) [9], Hierarchical Tree Alternative Path (HTAP) [5] and Heuristic Path Search (HPS) [4].

4.1 Performance Metrics

The analysis is calculated by separate calculations between the methods proposed and present. The dimensions include overhead control, energy usage, end-to-end latency and the distribution relationship for packets. The length of lifespan of the network is defined as the time in which the sensor network is broken into two or more networks. The control packets for path utilisation, implementation and maintenance

that specifically affect network lifetime are shared from the network layer point of view. This is overhead routing. A comprehensive simulation set based on the parameter is performed with the help of an NS2 simulator.

4.1.1 Average Control Packet Overhead

The sensor node sends the driving packets in the clustering region to construct and administer versatility. The total power consumption of various control packet protocols at varying sinking rate as shown in Fig. 3. As the figure shows, the control package overhead is slightly less in the planned HPS relative to the other protocols. An inline node in HTAP protects the information from the reference node. The information is sent to the pool when the request is sent to this inline node. There is an overhead control packet that floods the sink into the grouping area. The AACC protocol gives a special approach to planning trains and constructing stations. Nevertheless for metadata processing on a station and recovery from a drain, a controlling packet exchange is required. Both ring nodes preserve the position of the sink in HMOO. This makes it easy to extract the sink position. However as a result of the network service, the ring must be repaired by swapping control packets. This raises the ring length and the distance from the source or sink contributes to a higher energy consumption. The suggested BSO–PSO shall maintain the tree in the clustering field for data transmission. The controller packets are important to set the connection to the sink location. However, overhead power would be minimised in the proposed HPS. This is because the average distance between the cluster regions from the source or sink is lower than the other protocols.

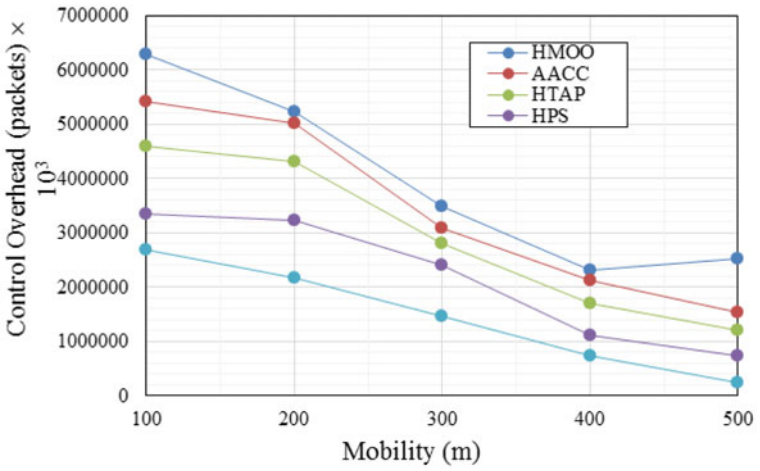


Fig. 3 Average control overhead

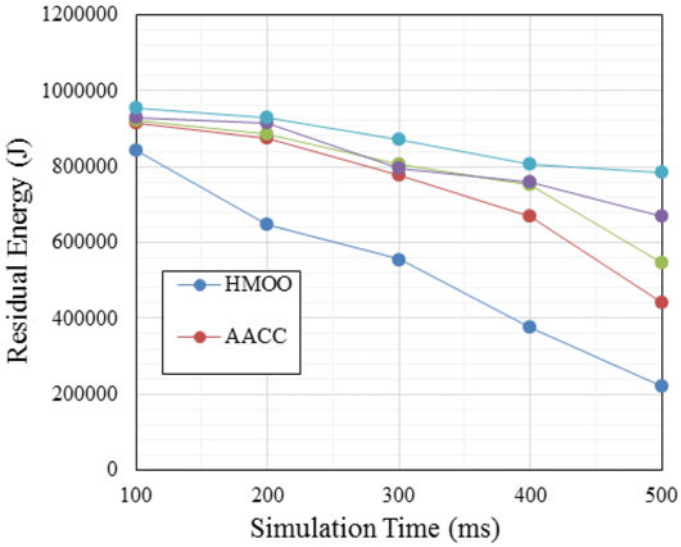


Fig. 4 Average energy consumption

4.1.2 Average Energy Consumption

Figure 4 demonstrates each node's overall energy usage of the numerous protocols. The increased overhead management kit has shown that the energy consumption of HTAP is greatest. The source node data is processed and the sink request is clustered in the region.

As the rate of fall increases, the HTAP's power usage can rise single-handedly. The proposed BSO–PSO does not need a drain, but has a mean distance longer than the proposed AACC, HMOO and HPS. This also reduces the overall energy consumption and relies on the lowered pace. The average distance between the origin and the expected HPS sink is approximately the same as the AACC and HMOO. However with the lower overhead command kit, the latest HPS improves existing protocols.

4.1.3 Average End-to-End Latency

Figure 5 displays the mean latency of various protocols at varying drop rates. It depends on the time it takes for the sink to be located and the information distributed to the sink. The recommended BSO–PSO transmits information directly to the backend tree. The tree transfers the data to the sink, because the sink is still connected. As a result, the lag is at the end of the day considerably lower. However when the HTAP module is reached, the inline node transmits the data. The suggested HPS needs less time to supply the data compared with AACC and HMOO. There is a shorter path between the clustering region and the root node.

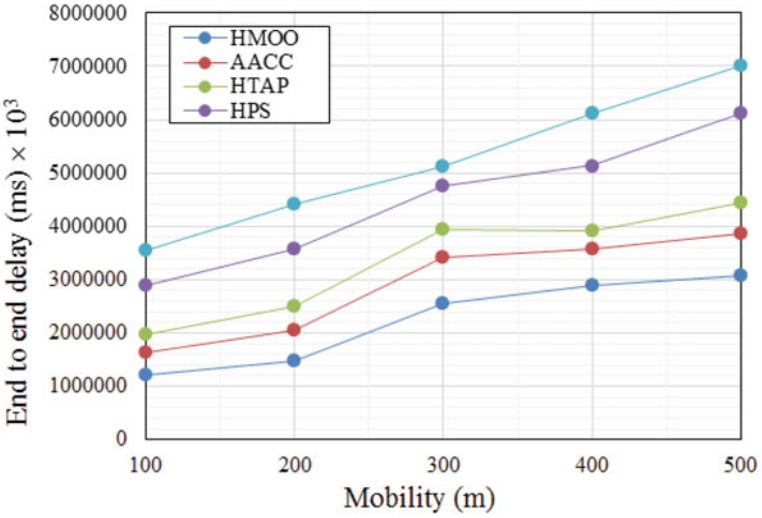


Fig. 5 Average end-to-end latency

4.1.4 Packet Delivery Ratio

Figure 6 indicates the packet distribution ratio of different protocols. The progress rate of reception of data on the sink is indicated. The proposed BSO–PSO retains

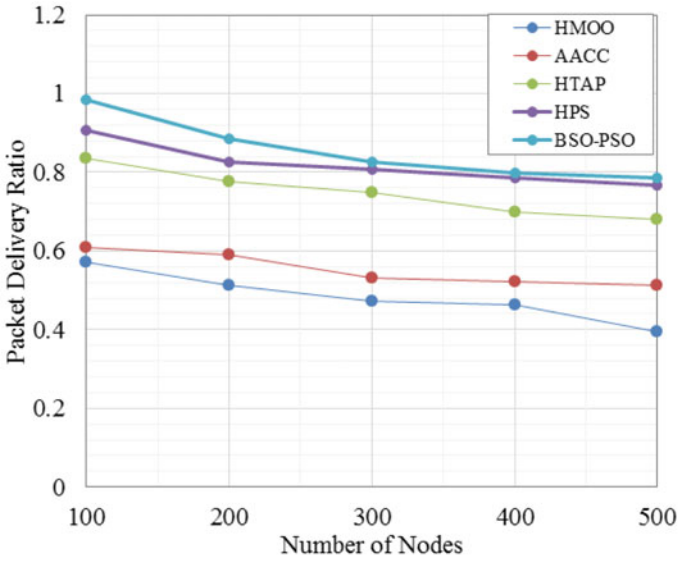


Fig. 6 Packet delivery ratio

the relationship between tree and drain. The transmission rate thus compares to that of other protocols. In HTAP, when the locale is reached, the data is saved and sent to the sink node. Consequently, data leakage is less common than most protocols. In contrast with the suggested HPS, the time required to hit the sink is higher in AACC and HMOO. The latency is minimised in data transfer. This would result in leakage of data from the sink to the new location.

5 Conclusion

This paper seeks to minimise energy usage and boost the stability of the network and the QoS. The suggested solution uses a modern congestion management based on prices to optimise the WSN network challenge. By optimising BSO optimization technologies, resources could be used. A QoS clustering over a hierarchical wireless sensor network providing the K-means algorithm. This technique achieves energy-efficient transmission. The results of the simulation showed that the new BSO–PSO performed the existing protocol for the purposes of end-to-end latency and transmission ratio. HPS is significantly less than the actual BSO–PSO energy consumption. In future work, the proposed protocol can be extended further to enhance energy efficiency in MAC layer with the routing protocol. Adopting the concepts of improved fault tolerance, safety and mobility in communication will be performed. Some error such as network coding or a forward error fix can be used to recover packet lost due to a failure in multipath routing.

References

1. Ghaffari A (2015) Congestion control mechanisms in wireless sensor networks: a survey. J Netw Comput Appl 52:101–115
2. Monowar MM, Rahman MO, Pathan ASK, Hong CS (2012) Prioritized heterogeneous traffic-oriented congestion control protocol for WSNs. Int Arab J Inf Technol 9(1):39–48
3. Yuvaraj N, Karthikeyan T, Praghash K (2020) an improved task allocation scheme in serverless computing using gray wolf optimization (GWO) based reinforcement learning (RIL) approach. Wireless Personal Communhttps://doi.org/10.1007/s11277-020-07981-0
4. Sangeetha G, Vijayalakshmi M, Ganapathy S, Kannan A (2018) A heuristic path search for congestion control in WSN. In: Industry interactive innovations in science, engineering and technology. Springer, Singapore, pp 485–495
5. Sergiou C, Vassiliou V, Paphitis A (2013) Hierarchical tree alternative path (HTAP) algorithm for congestion control in wireless sensor networks. Ad Hoc Netw 11(1):257–272
6. Sivaram M, Porkodi V, Mohammed AS, Manikandan V (2019) Retransmission DBTMA protocol with fast retransmission strategy to improve the performance of MANETs. IEEE Access 7:85098–85109. https://doi.org/10.1109/ACCESS.2019.2918723
7. Yuan H, Yugang N, Fenghao G (2014) Congestion control for wireless sensor networks: a survey. In: The 26th Chinese control and decision conference (2014 CCDC). IEEE, pp 4853–4858

8. Kousik NV, Raja RA, Saravanan M (2020) Automatic skull-face overlay and mandible articulation in data science by AIRS-Genetic algorithm. Int J Intell Netw 1:9–16
9. Zhuang Y, Yu L, Shen H, Kolodzey W, Iri N, Caulfield G, He S (2018) Data collection with accuracy-aware congestion control in sensor networks. IEEE Trans Mob Comput 18(5):1068–1082
10. Singh K, Singh K, Aziz A (2018) Congestion control in wireless sensor networks by hybrid multi-objective optimization algorithm. Comput Netw 138:90–107
11. Wang C, Sohraby K, Lawrence V, Li B, Hu Y (2006) Priority-based congestion control in wireless sensor networks. Trustworthy Comput 1
12. Md. Rahman, Monowar MM, Hong CS (2008) A QoS adaptive congestion control in wireless sensor network international conference on advanced communication technology
13. Monowar MM, Rahman MO, Pathan A-SK, Hong CS (2008) Congestion control protocol for wireless sensor networks handling prioritized heterogeneous traffic. International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services, Dublin, Ireland
14. Yaghmaee MH, Adjeroh DA (2009) Priority-based rate control for service differentiation and congestion control in wireless multimedia sensor networks. Comput Netw 53(11):1798–1811
15. Tao L, Yu F (2010) ECODA: enhanced congestion detection and avoidance for multiple class of traffic in sensor networks. IEEE Trans Consu Electron
16. Sridevi S, Usha M, Lithurin (2012) Priority based congestion control for heterogeneous traffic in multipath wireless sensor networks ICCCI
17. Sun Y, Bian J, Tang L (2012) Priority node-based congestion control algorithm for wireless sensor networks WiCOM

Financial Technology Development Around the World and Lessons for Emerging Markets

Toan Linh Vu Le and Duy Khanh Pham

Abstract Fintech has always been a "hot" keyword for the banking and finance sector in the world. The US and China are the world leaders in fintech with a complete fintech ecosystem. Singapore and Indonesia are the two leading countries in Fintech in Southeast Asia. Looking at the current situation of the Fintech market in other countries and comparing it with the situation in Vietnam, it is possible to draw lessons from experience and make suggestions for parties to develop the Vietnamese fintech market.

Keywords Fintech · Blockchain · Startup

1 Introduction

Advantages for Fintech in Vietnam

The Internet growth rate in Vietnam is higher than the regional average and is gradually setting foot in the top of the world's leading countries in this field. As of January 2021, the proportion of mobile subscribers using smartphones and tablets accounted for 96.9 and 31.9%, respectively, is higher than the world average. Vietnam's Internet penetration rate has reached 70% of the population, with a population of 68.72 million Internet users. Banking-finance applications account for 40.1% of popular applications installed on smartphones [11].

T. L. Vu Le
Faculty of Banking and Finance, Van Lang University, Ho Chi Minh City, Vietnam
e-mail: toan.lvl@vlu.edu.vn

D. K. Pham (✉)
School of Banking, University of Economics Ho Chi Minh City (UEH), Ho Chi Minh City, Vietnam
e-mail: duy.pham@ueh.edu.vn

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_23

Challenges for Fintech in Vietnam

Mechanisms, policies, and legal frameworks are incomplete and synchronous for fintech activities. The unclear law makes businesses in the industry extremely cautious. Mr. Nguyen Ba Diep, Deputy General Director of MoMo company, said that current regulations force users to have a bank account—which causes difficulties. There is no small obstacle to the development of fintech because this new field was created to ensure that everyone who is "forgotten" by banks also has access to finance.

Cryptocurrencies such as Bitcoin are not considered a legal means of payment in Vietnam, nor are they recognized as property in the Vietnamese legal system. The Ministry of Industry and Trade does not currently recognize Bitcoin as a good or service and denies its authority over Bitcoin, abolishing the basis for tax collection. Bitcoin exchanges are still taking place but are "not protected by law."

Technology infrastructure, the financial market has not met the requirements. Vietnam's security technology still has many limitations. The development speed does not keep up with the scale development of fintech [12].

2 Methodology

Authors mainly use qualitative analysis based on descriptive statistics and charts related to fintech development in USA, China, Singapore, and Vietnam.

3 Main Findings

Fintech development around the world

The United States of America

Currently, the US is considered the largest and oldest fintech hub with the resources to establish some of the world's leading fintech ecosystems from the presence of the world's top universities, great tech talent pool for financial expertise. The leading US fintech cities, the world's leading fintech hubs, are San Francisco and Silicon Valley. According to CB insights, some of the world's biggest fintech companies are from here, including Stripe, specializing in online payment processing for small businesses, the most valuable fintech company at $95 billion, Chime ($14.5 billion), and Plaid Technologies ($13.5 billion). According to the Startup Genome report, Silicon Valley ranks first in most of the startup sub-sector rankings from AI, robotics, to life sciences and is the main unicorn factory in the world, producing more than 96 to December 2019. The top seven AI investors globally, including Google, Facebook, and Apple, are based here [4].

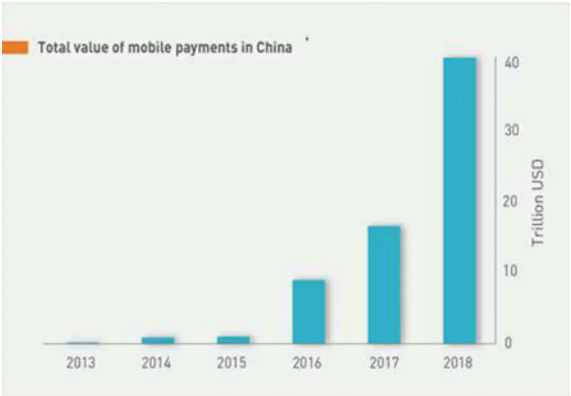


Fig. 1 Total value of mobile payments in China *Source* Financial Times

China

Among three areas of fintech growth, the first is mobile payments [17]. The number shows that China experienced about 50% of the world's digital payments in 2017. Cashless transactions take into account of one-third of consumer payments, while comparing to a quarter of US users, three-quarters of users' Chinese smartphone performed mobile point-of-sale purchase in that year. Alipay, the payments division of Alibaba, has quickly become the mobile wallet of choice for many. Tencent also introduced a payment function in its popular mobile app WeChat, which has 500 million users. AliPay and WeChatPay control 93% of the mobile payment segment. Mobile payments with QR codes on the WeChat and Alipay reached 38 trillion yuan ($5.5 trillion) in 2016, more than five times the size of the US market [3]. Figure 1 shows the total value of mobile payments in China in 2018 reached about 40 trillion USD, double that of 2017, an outstanding growth.

Next, China developed online loans which takes three-quarters of the world market. For instance, Alibaba and JD.com, China's two largest e-commerce portals, easily lend small amounts.

The third area is an investment. In the past, Chinese savers could only choose to deposit money in a bank with a low but safe interest rate or pour money into the stock market; the return can be higher, but it is risky. "There is no middle ground," said Huang Hao, vice president of Ant Financial. Moreover, that gap has been opened by fintech. In 2013, it experienced the hottest that is Alibaba's launch of the Yu'e Bao online fund. In 2014, Tencent launched Licaitong, an online fund platform that connects to WeChat [17].

Currently, blockchain and cryptocurrency are two prominent fintech sectors in Singapore. In Singapore, the nation experienced 19% (among 430 startups fintech) is active in blockchain and crypto, making it the largest segment, pre-payments with 16%, investment and financial technology with 14%, and regtech with 11% (see Fig. 2).

Singapore also developed in fintech very much, for instance, a project to link PayNow system of the nation with PayPromt system Thailand in order to facilitate

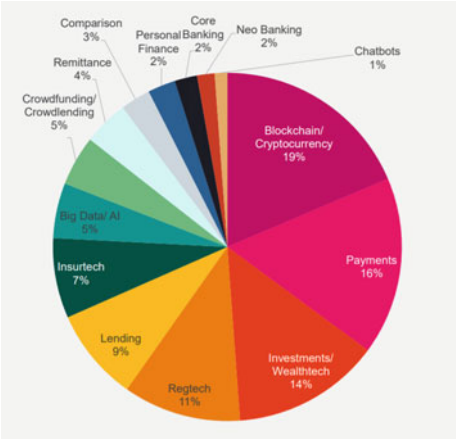


Fig. 2 Breakdown of the Singapore fintech ecosystem *Source* Singapore Fintech Report 2021, Fintech News Singapore

cross-border payments cheaper and faster. There are more than 1000 startups (in fintech field—online lending and personal finance) in Singapore the island with 5.7 m people living, and the country is trying to become a hub fintech [6].

4 Discussion and Conclusion

Fintech development in Vietnam

Fintech activities in Vietnam also developed quickly, and it ranks just after Singapore and as the 2nd in ASEAN (as with 51% fintech funding, accounting 36% of all of the region's fintech investment) [7].

Figure 3 shows Vietnam currently has about 123 fintech startups coming from many segments such as digital payments, alternative finance, wealth management, blockchain. During 2017–2020, the country has experienced an increasing ratio up to 180% for fintech startups.

Payment with 31% of all fintech startups remains the largest segment. Second place is P2P lending with 17%. The top five sectors, included Payments, P2P lending, blockchain, POS, wealth management, have taken 76% of the total number (Fig. 4). Vietnam has 39 licensed non-bank payment service providers as of October 2020. Fintech companies in the payment segment attract up to 98% of the total funding for the Vietnamese fintech market. Momo, Payoo, Moca, ZaloPay, VietelPay are the most commonly used e-wallets in Vietnam today.

According to statistics from the State Bank of Vietnam, as of 2019, out of a total of nearly 100 million people, 4.2 million people use e-wallets. The above figure shows that payment is a highly competitive field with many development opportunities.

P2P and blockchain are two areas that attract great interest from investors. The number of startups in these two segments increased from 2017 (less than 5 firms) to

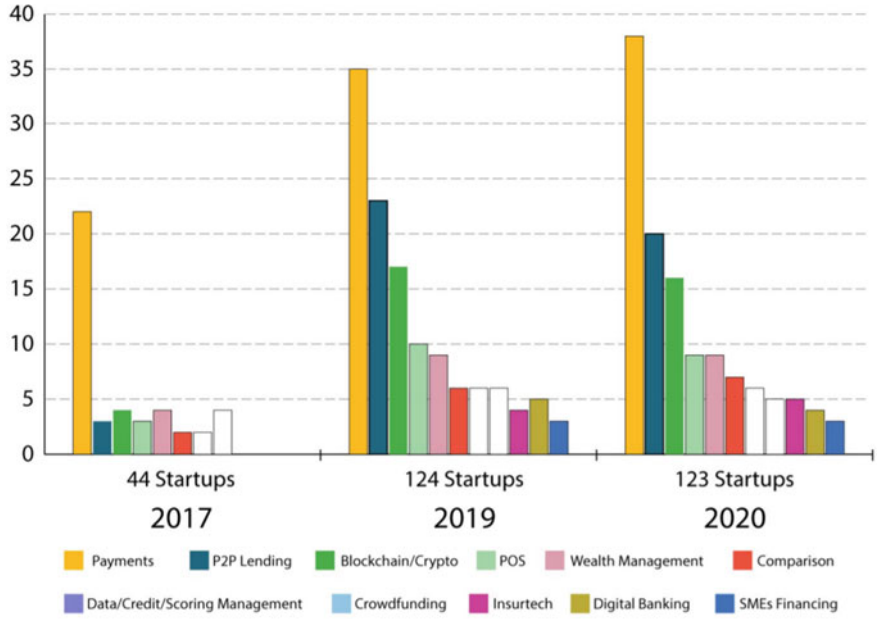


Fig. 3 Situation of fintech over years *Source* Report 2020-Vietnam Fintech, Fintech News Singapore

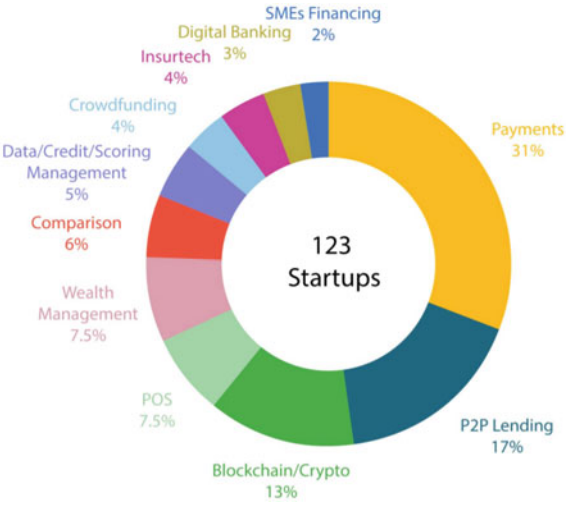


Fig. 4 Vietnam's fintech players in 2020 *Source* Vietnam Fintech Report 2020, Fintech News Singapore

2020 (over 15 firms). In the last 3 years, the nation experienced the digital banking and Insurtech emergence, as well as SMEs financing.

Proposals for banks

Banks are suggested to cooperate with fintech companies regarding information exchange and technological advances to take advantage of the technology infrastructure provided by fintech companies instead of spending time and money to build it. Cooperation with fintech companies supports banks to expand, diversify and enhance products and services. Partnerships with universities in banking, finance, and information technology help banks recruit talented personnel. Last but not least, banks are to organize training programs to provide employees with new technical knowledge [11].

References

1. Anh H (2020) *ictnews.vietnamnet.vn.* Retrieved August 15, 2021, from ictnews.vietnamnet.vn: https://ictnews.vietnamnet.vn/cuoc-song-so/my-giam-ganh-nang-phap-ly-cho-cac-fintech-264187.html
2. CB Insights (2021) www.cbinsights.com. Retrieved September 3, 2021, from https://www.cbinsights.com/research/fintech-startups-us-map/
3. CGAP (2019) www.cgap.org. Retrieved September 3, 2021, from https://www.cgap.org/research/publication/china-digital-payments-revolution
4. Fintech News America (2021) *fintechnews.am.* Retrieved September 3, 2021, from https://fintechnews.am/fintech-usa/46215/top-5-fintech-cities-in-the-usa/
5. Fintechnews Indonesia (2021) Indonesia Fintech Report and Map 2020. Fintechnews Singapore
6. Fintechnews Singapore (2021). Singapore Fintech Report 2021: Blockchain Dominates Singapore's Fintech Scene. Fintechnews Singapore
7. Fintechnews Vietnam (2020) 2020 Fintech Vietnam Report and Startup Map: Fintech Startups Tripled since 2017. Fintechnews Singapore
8. Hoang YH, Nguyen NT, Nguyen DT, Vu NB, Tran LH (2021) *diendandoanhnghiep.vn.* Retrieved September 3, 2021, from
9. https://diendandoanhnghiep.vn/ngan-hang-bat-tay-fintech-ky-3-cac-giai-phap-ho-tro-206091.htm
10. Huynh TT (2019) thitruongtaichinhtiente.vn. Retrieved September 3, 2021, from https://thitruongtaichinhtiente.vn/fintech-va-su-doi-moi-cac-san-pham-dich-vu-tai-chinh-ngan-hang-25313.html
11. Lao D (2021) specials.laodong.vn. Retrieved September 3, 2021, from https://specials.laodong.vn/chuyen-doi-so-tai-viet-nam-va-nhung-thong-ke-an-tuong-2021/
12. Le NH (2018) sob.ueh.edu.vn. Retrieved August 1, 2021, from http://sob.ueh.edu.vn/wp-content/uploads/2018/11/16.-LE-HUYEN-NGOC.pdf
13. Mordor Intelligence (2020) www.mordorintelligence.com. Mordor Intelligence. Retrieved from https://www.mordorintelligence.com/industry-reports/china-fintech-market
14. Nguyen YH (2020) Fintech in Vietnam and its regulatory approach. In: Fenwick M, Van Uytsel S, Ying B (eds), Regulating fintech in Asia. perspectives in law, business and innovation. Springer, Singapore, pp 115–138
15. Phong D (2020) viettimes.vn. Retrieved August 15, 2021, from https://viettimes.vn/bai-4-phap-ly-ve-fintech-cang-cham-cang-lo-post132369.html

16. Van Phong (2020) thesaigontimes.vn. Retrieved August 15, 2021, from https://www.thesaigontimes.vn/308692/giai-bai-toan-nhan-luc-lap-khoang-trong-chinh-sach-cho-fintech.html
17. Van Quoc (2017). nhipcaudautu.vn. Retrieved August 15, 2021, from https://nhipcaudautu.vn/cong-nghe/nho-dau-ma-trung-quoc-di-truoc-ca-the-gioi-ve-fintech-3318134/

Performance Comparison of Brain Tumor Segmentation Algorithms

Divya Mohan, V. Ulagamuthalvi, and Nisha Joseph

Abstract In the medical field history, the phenomenon of image segmentation plays an important role. The proper segmentation algorithm will help very much to reduce the time taken for diagnosis. This paper intends to compare the performance of various algorithms used to segment tumor-affected voxels present in Magnetic Resonance Imaging (MRI) of brain tumors. As a brain tumor is a dangerous and life-threatening disease that may affect persons in any age span, early diagnosis will help clinicians make the correct decisions regarding the treatment method and save a precious life. The automation of brain tumor detection from an MR image is very much important. The MR images give good detail of brain tumors compared to other medical imaging modalities. Comparison of different algorithms implemented to separate brain tumor cells from normal brain cells using different performance metrics. The performance parameters under consideration are accuracy, precision, recall, sensitivity, specificity, F1 score, Dice similarity coefficients. The paper concludes that the algorithm Fuzzy C Means combined with Graph cut provided 99.66% of acuracy on the dataset ISBR compared with other algorithms. So it is able to say that Fuzzy C means with Graph Cut is the best out of all algorithms under consideration.

Keywords Magnetic resonance imaging · Brain tumor · Performance metrics · Segmentation

D. Mohan (✉) · V. Ulagamuthalvi
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India
e-mail: divya.mohan@saintgits.org

V. Ulagamuthalvi
e-mail: ulagamuthalvi.cse@sathyabama.ac.in

N. Joseph
Department of Computer Science and Engineering, SAINTGITS College of Engineering (Autonomous), Kerala, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_24

1 Introduction

Cancer is a life-threatening disease these days and can appear anywhere in the human body. The brain tumor is especially very dangerous as it may appear in any age group of humans. There are different kinds of brain tumors:-Benign and Malignant tumors. World Health Organization (WHO) categorized tumors into four grades. Grade I and II are Benign types with a standardized structure, low risk as it does not produce cancer cells and is very tedious to remove from normal brain tissues through surgery or radiation. Grade III and IV are malignant tumors which are non-uniform, high-risk tumors since it contains cancer cells.

There exist different modalities for medical imaging like X-Ray, CT-Scan, and MRI. There exist different MRI methods to better the information taken from MR images. These are named T1, T2, T1c, T2flair. They are as follows T1-weighted spin–lattice relaxation (T1), T2-weighted spin–spin relaxation (T2), T1-weighted MRI with contrast improvement (T1c), and T2-weighted MRI with fluid attenuation inversion recovery (T2flair), where each modality provides different kinds of data regarding tumor pixels [6] (Fig. 1).

Because of the patients' less exposure to the ionized environment like radiations, MR Imaging is more likely advised. Also, MRI is providing more detail of the tumor [10].

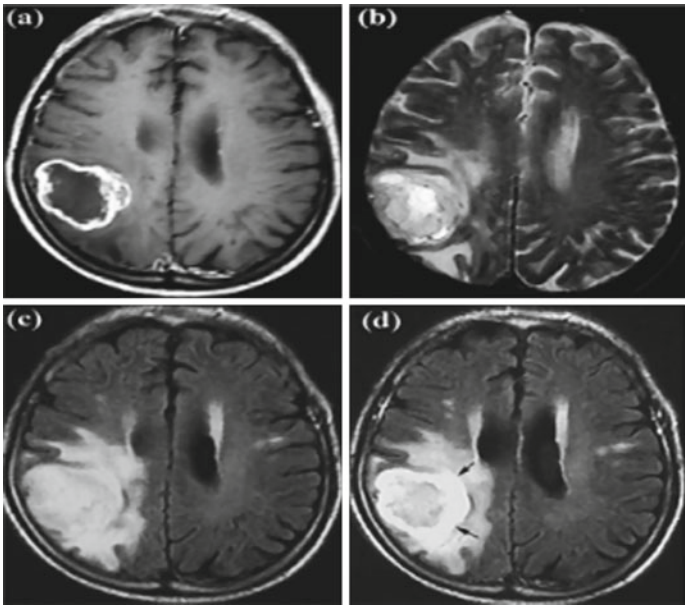


Fig. 1 **a** T1 weighted MRI: **b** T2 weighted MRI: **c** FLAIR d FLAIR with contrast enhancement [10]

The primary process involved in automated tumor separation from normal tissues is as follows. The first step is to preprocess captured MR images like skull striping, patching, denoising, etc. The second step is applying the image to segmentation to separate tumors from normal tissues.

Various performance parameters selected to compare different segmentation algorithms' efficiency are accuracy, sensitivity, and specificity.

Accuracy quantifies the efficiency of the segmentation algorithm [1].

It is calculated by

$$\text{Accuracy} = \left(\frac{\text{TP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}\right) * 100 \tag{1}$$

TP is calculated by the number of samples properly divided using the process from the normal brain cells, TN is the number of samples correctly separated by the process as normal brain tissues, FP is the number of samples separated by the tumor. In contrast, the samples are standard, and FN is the number of samples incorrectly separated by the process as usual.

Sensitivity represents how well an algorithm segment the tumor region from a specific brain region.

$$\text{Sensitivity} = \left(\frac{\text{TP}}{\text{TP} + \text{FN}}\right) * 100 \tag{2}$$

Specificity denotes the performance of an algorithm in segmenting normal tissue in an MR image.

$$\text{Specificity} = \left(\frac{\text{TN}}{\text{TN} + \text{FP}}\right) * 100 \tag{3}$$

$$\text{Dice Score} = \left(\frac{2 * \text{TP}}{2 * \text{TP} + \text{FN} + \text{FP}}\right) * 100 \tag{4}$$

Positive Predictive Value [2] is the ratio of no samples correctly segmented as tumor cells on a total number of samples segmented as tumor cells.

$$\text{PPV} = \left(\frac{\text{TP}}{\text{TP} + \text{FP}}\right) * 100 \tag{5}$$

$$\text{Jackard Index} = \left(\frac{\text{TP}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}\right) * 100 \tag{6}$$

$$\text{Recall} = \left(\frac{\text{TP}}{\text{TP} + \text{FN}}\right) * 100 \tag{7}$$

2 Analysis of Various Segmentation Techniques

A. ***Random Forest classifier for segmentation***

This work utilized four sequences of MR Images T1, T1c, T2, Flair as input images. Range of Histogram, Center Symmetric Local Binary Pattern (CSLBP), and Mean/Average is extracted from uniting all the four sequences. In the Segmentation phase, the Random Forest (RF) classifier is used to separate all the tumor-affected areas [2].

B. ***Brain Tumor Segmentation Using K-Means Clustering and Fuzzy C-Means Algorithms***

This work was implemented in the segmentation phase using two different simple clustering algorithms –K Means and Fuzzy C Means. The clustering process is an unsupervised task. The K Means algorithm's segmentation accuracy was 99%, whereas FCM was 95% [3].

C. ***Segmenting brain tumor using the modified watershed algorithm in MRI image***

This work segmented the 2D MR images of brain tumors into two steps. First edge detection was performed using a canny edge detector, and then segmentation was implemented using a modified watershed algorithm to get better performance over the watershed algorithm [4].

D. ***Brain Tumor Segmentation in MR Images Using Deep Learning***

proposed a hybrid Convolutional Neural Network containing nine convolutional layers that combined two path and three path networks to extract local and contextual features in the input. The computation time is significantly less than individual two paths and three path convolutional neural Networks [5].

E. ***Brain Tumor Segmentation Based on Hybrid Clustering and Morphological Operations***

As far as K-means++ clustering is concerned, it is employed with the Gaussian kernel-based fuzzy C-means algorithm for segmenting images to improve the algorithm's stability by reducing the clustering parameters' sensitivity [7].

In cs-FCM, the idea is to apply an accommodating result to the auxiliary variables correlating to each pixel to reduce the algorithm's reactivity to noise and intensity diverseness concerning MRI data [8].

In spatial information for infrared ship segmentation (sFCM), improved the FCM in two areas (1) general spatial accession information formulated on ship targets (2); refining of the local space curtailment via the Markov random field using the of the ship's target contour spatial shape information[9].

F. ***Hybrid fuzzy brain-storm optimization algorithm (FBSO)***

The average brain storm optimization was using k means clustering. But the proposed algorithm combined with fuzzy c means clustering with a Brainstorm optimization algorithm to get robust segmentation results [10].

G. ***Modified morphological-based Fuzzy-C-Means (M- FCM) algorithm (GAPBFC M)***

This work implemented M-FCM to separate cancer cells from noncancerous brain cells. The membership function value of each cluster point was revised to implement M-FCM [11].

H. ***hybrid model of Fuzzy C-Means (FCM) clustering integrated with the Genetic Algorithm (GAPBFC M)***

This work combined initial genetic algorithm stages with probability-based fuzzy c Means to eliminate over-segmentation [12].

I. ***Fuzzy C Means with Graph Cut***

Grouping Fuzzy C Means and Graph cut method eliminates cluster center initialization highlight edges and provides faster segmentation [1].

3 Result and Discussion

Most of the algorithms were tested using the BRATS dataset. Table 1 brief the performance of all the above-discussed algorithms in terms of various performance metrics.

Table 1 Performance of segmentation algorithms in terms of various performance metrics (approx)

Algorithm	Accuracy/ DiceScore	Data set	Specificity/Jackard Index/PPV	Sensitivity
KFCM	96.66%	IBSR	96%	95.5
Fuzzy C means with graph cut	99.66%	IBSR	94%	96.5
FCM	92.4%	IBSR	90%	90.5
Random forest classifier	89(DiceScore)	BRATS 2015	90(PPV)	92
Modified watershed algorithm with FPGA	98.81(Modified watershed algorithm)	Private data	84.9	99.39
	96.72(Watershed)	Private data	80.1	95.37
Deep learning	86(DiceScore)	BRATS 2013	91	86
GAPBFC M[12]	97.07/91.47 Dice Score	Private data	95.4(JackardIndex)	
M-FCM	97.89	Private data	100	97.75
FBSO	93.85	BRATS 2018	88.96	94.39
Hybrid clustering and morphological operations	90.45(s-FCM)	BRATS 2012	99.32	90.97
	88.9(Cs-FCM)	BRATS 2012	99.16	90.63
	92.56(K Means ++ and Gaussian kernel-based fuzzy C means clustering)	BRATS 2012	99.41	96.6

KFCM offered the accuracy 96.66% on the ISBR dataset. Fuzzy C means and Graph Cut had an accuracy 99.66% on the same dataset. FCM provided 92.4% on ISBR dataset. Random forest classifier offered 89% dice score. Modified watershed with FPGA provided 98.81% accuracy and watershed provided 96.72% accuracy on tested private dataset. The deep learning method provided 86% dice score on BRATS 2013 dataset. GAPBFCM guaranteed 97.07% accuracy on tested private data. M-FCM offered 97.89% accuracy on tested private data. FBSO provided 93.85% accuracy on BRATS 2018 dataset. Hybrid clustering and morphological operation gave 90.45%, 88.9% and 92.56% using s-FCM, CS-FCM and K-Means with Gaussian kernel based fuzzy C means respectively on dataset Brats 2012 data set. Analyzing various brain tumor segmentation algorithms using dataset IBSR, it is found that, Fuzzy C Means with Graph Cut Algorithm [1] shows the best performance in terms of accuracy with specificity 94% and sensitivity of 96.5%. While using BRATS Dataset, the FBSO algorithm was superior in terms of accuracy.

4 Conclusion

This paper evaluates the performance of different segmentation algorithms on MRI of brain tumors. This paper also studies the impact of various performance metrics to select an algorithm for brain tumor segmentation. Brain tumor separation from normal tissue has a vital role in the early stages of tumor diagnosis. For clinical purposes, the tumor separation from MR Images is done manually. This paper compares different segmentation techniques on MR Images of brain tumors to conclude with the best-implemented segmentation approach in terms of the listed performance metrics. In this comparison, Fuzzy C Means with Graph Cut algorithm [1] yields better accuracy under consideration. This paper only discussed segmentation techniques based on soft computing, machine learning and a hybrid of both. As a future enhancement step, more deep learning-based segmentation techniques can be considered for performance comparison using the same metrics.

References

1. Mathew ST (2020) A semi-automatic brain tumor. J Critical Rev 7.8:2213–2222
2. Nisha J, Murugan D, Thomas BJ (2020) Patient-specific brain tumor segmentation using context sensitive feature extraction in MR images. Int J Comput Digital Syst 9.6:1091–1097.https://doi.org/10.12785/ijcds/090607
3. Rushikesh K et al (2018) Brain tumor segmentation using K-means clustering and fuzzy C-means algorithms and its area calculation and disease prediction. https://doi.org/10.15680/IJIRSET.2018.0706020
4. Sivakumar V, Janakiraman N (2020) A novel method for segmenting brain tumor using modified watershed algorithm in MRI image with FPGA. Biosystems 198:104226. https://doi.org/10.1016/j.biosystems.2020.104226

5. Sajid S, Hussain S, Sarwar A (2019) Brain tumor detection and segmentation in MR images using deep learning. Arab J Sci Eng 44(11):9249–9261. https://doi.org/10.1007/s13369-019-03967-8
6. Solmaz A, Pour FT (2014) A hybrid approach for detection of brain tumor in MRI images. In: 21th Iranian conference on biomedical engineering (ICBME). IEEE
7. Chong Z et al (2019) Brain tumor segmentation based on hybrid clustering and morphological operations. Int J Biomed Imaging. https://doi.org/10.1155/2019/7305832
8. Kumar AS et al (2015) Conditional spatial fuzzy C-means clustering algorithm for segmentation of MRI images. Appl Soft Comput 34:758–769
9. Xiangzhi B et al (2015) Infrared ship target segmentation based on spatial information improved FCM. IEEE Trans Cybernet 46.12:3259–3271
10. Narmatha C et al (2020) A hybrid fuzzy brainstorm optimization algorithm for the classification of brain tumor MRI images. J Ambient Intell Humanized Comput 1–9. https://doi.org/10.1007/s12652-020-02470-5
11. Tessy G et al (2020) A professional analysis on the segmentation of brain cancer in MR images using M-FCM. In: International conference on inventive computation technologies (ICICT). IEEE
12. Latha C, Perumal DRK (2020) Segmentation of brain tumor MR images in soft computing techniques. J Theoret Appl Info Technol 98.19

AROSTEV: A Unified Framework to Enhance Secure Routing in IoT Environment

R. Stephen, A. Cecil Donald, A. Dalvin Vinoth Kumar, B. J. Hubert Shanthan, and L. Arockiam

Abstract The Internet of Things (IoT) is a global network which collects, process, and analyzes the data. IoT sensors and devices are limited to low memory, power, and processing capabilities. The RPL is a proactive routing protocol which is mainly intended for the IoT. There is a possibility of routing vulnerabilities, which masquerade the data in IoT environment. In order to overcome this problem, a framework called AROSTEV is proposed which comprises of three techniques such as RDAID, RIAIDRPL, and E2V. The primary objective of AROSTEV framework is to detect and mitigate the routing attacks such as rank decreased attack (RDA), rank increased attack (RIA), and rank inconsistency attack (RInA), respectively. Each technique takes the responsibility to progress its activity against the internal routing attacks. This framework can be used to implement the smart city environment.

Keywords Internet of things · Routing attacks · Rank inconsistency · RPL · AROSTEV

R. Stephen (✉) · B. J. H. Shanthan
Kristu Jayanti College (Autonomous), Bengaluru 560077, India
e-mail: r.stephen@kristujayanti.com

B. J. H. Shanthan
e-mail: Hubert@kristujayanti.com

A. C. Donald
Assistant Professor, Department of Computer Science, CHRIST (Deemed to be University), Central Campus, Bangalore 560029, India
e-mail: cecil.donald@christuniversity.in

A. D. Vinoth Kumar
Assistant Professor, Department of Computer Applications, School of Engineering, Dayananda Sagar University, Bangalore 560068, India
e-mail: dal_win@ymail.com

L. Arockiam
St. Joseph's College (Autonomous), Tiruchirappalli 620002, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_25

1 Introduction

The Internet of Things (IoT) is a global network which collects, process, and analyzes the data. The IoT technologies are used in various application domains. But, deployment of IoT applications is a challenge in the context of security issues. So, the different types of security threats tend to rise among IoT applications. Thus, providing a secure routing among IoT sensor nodes is a fundamental factor to be considered in the IoT environment. IoT has self-configuring and unique characteristics which require security from different routing attacks. The IoT networks employ IP-based low power and lossy networks (LLNs) in which heterogeneous devices are connected. For example, the sensors gather data from the environment (temperature, weather condition), hospital management (patient's details), and exchange this sensitive data between devices. During data transmission between IoT sensor nodes, this sensitive data may be hacked and become invalid due to the network routing attacks. Hence, the IoT systems need a secure routing framework to protect the sensor data.

Many security mechanisms were proposed by various researchers. These mechanisms hold important space in protecting the protocol from attacks like man-in-the-middle and denial of service. The SVELTE [1] is a real-time intrusion detection system (IDS) for secure routing in IoT. But, this IDS module was not deployed into any application domains. The INTI [2] is an IDS module against the sinkhole attacks. The INTI only considered the sinkhole attacks, but not the other attacks. This module was not implemented in any applications. The InDRes module was proposed by the authors [3] for detecting sinkhole attacks and not implemented in any applications.

The technique, called RDAID [4], is proposed to identify and mitigate the rank decreased attack which is illegally decreased its rank value. The RDAID is also evaluated with various parameters such as detection accuracy, control message, delivery ratio, and packet loss. The author in [5] proposed a technique called RIAIDRPL against rank increased attack which creates a loop in the RPL network. The author in [6] proposed architecture called E2V that consists of three phases. DvirAmit et al. [7] prevented an internal attacker to publish decreased rank value. GuoJianlin et al. [8] provided an innovative rank computation scheme and a loop-free local route repair mechanism which eliminated routing loops in RPL. SehgalAnuj et al. [9] proposed the approach based on a fixed threshold against DODAG inconsistency attacks such as DoS and black hole attacks. This paper presented a mitigation approach that allowed nodes to dynamically adjust adjacent to DODAG inconsistency attack. PerreyHeiner et al. [11] focused on routing security of RPL. The authors proposed authentication security service to RPL topology called trust anchor interconnection loop (TRAIL).

From the literature, it is observed that a secured framework for IoT environment is the need of hour. Many researchers proposed the security mechanisms, but they have not provided a full-fledged framework. This proposed framework AROSTEV consists of three different techniques for protecting the data in IoT environment. These techniques are proposed to evaluate rank consistency for RPL protocol, where rank and energy (RE) play a vital role. The proposed techniques RDAID, RIAIDRPL,

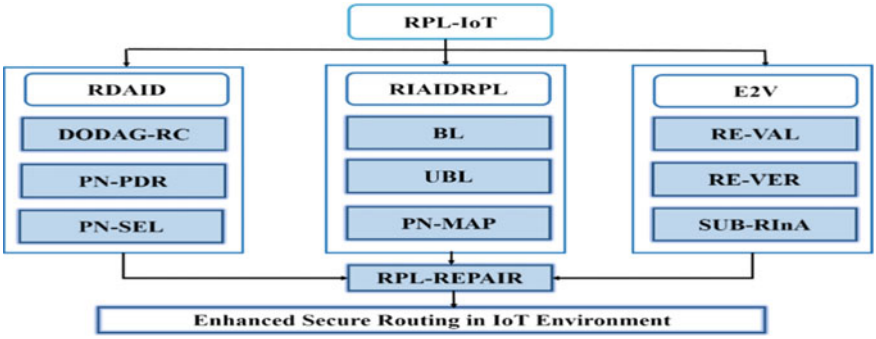


Fig. 1 Structure of the AROSTEV framework for IoT framework

and E2V are integrated into a novel framework. The aim is to offer secure routing for IoT data and also to avoid loops in the network topology.

2 Structure of AROSTEV Framework

The AROSTEV framework consists of three techniques which perform rank consistency in RPL-based IoT environment to enhance secured routing. The three techniques are RDAID, RDAIDRPL, and E2V are developed for different scenarios. The techniques are proposed as part of the RPL protocol for IoT (RPL-IoT).

The structure of the framework AROSTEV and components is shown in Fig. 1. Each proposed technique has three procedures, and each procedure is developed for a specific purpose. The aim is to offer secure routing in the IoT network.

3 Components of the AROSTEV Framework

The three techniques are incorporated into a framework called AROSTEV. Each technique has three procedures which are considered as the components of the AROSTEV framework. The components are listed below:

3.1 RDAID

The first component of the framework is called as rank decreased attack identification (RDAID). The RPL allows nodes to self-configure themselves, in which the RDA attacker node changes the rank by decreasing the rank value and creates more traffic in the network. RDAID consists of the following three procedures:

DODAG-RC. DODAG rank calculation (DODAG-RC) procedure is executed from border router (Root). After rank calculation, the source node selects a preferred parent node (Intermediate node) to route the data packets.

PN-PDR. During the parent selection, the child node considers the parent node's packet delivery ratio (PN-PDR) to send DIO_ACK (Acknowledgment message).

PN-SEL. This procedure parent node selection (PN-SEL) is executed from existing child nodes. The RDAID technique is deployed in the source nodes. These nodes scrutinize the PDR ratio before selecting a preferred parent node. The PDR ratio is calculated for each transmission.

3.2 RIAIDRPL

The second component of the framework is called as RIAIDRPL which detects RIA and performs secured routing by avoiding RPL loop. In RPL DODAG, the intruder illegitimately increases the rank value and creates a loop between parent and child nodes. So, it is important to detect and mitigate the loop in the DODAG. The RIAIDRPL traverses through each node to find the loop occurrence. It consists of the following three procedures:

BL. Before traversing the RIAIDRPL, it places all the nodes in the blocked list (BL), which represents the unvisited nodes. This procedure is deployed in the root node (BR) to handle this list.

UBL. The RIAIDRPL places the visited nodes in the Un_Blocked List (UBL) and also deletes the visited nodes from the blocked list. This procedure is also deployed in the root node.

PN-MAP. This procedure is done at the root node. The parent mapping (PN-MAP) is an important process than others. This process is done simultaneously with the UBL procedure. During the RIAIDRPL traversing, it analyzes the node's parent list. If any parent node's DIO message is inconsistent, it creates loop and considered to be a RIA type of attacker node. The proposed RIAIDRPL technique detects the RIA attacker node by using PN-MAP.

3.3 E2V

The third component is E2V technique which is deployed in the root node. The E2V receives all the nodes updated DIO message and ensures the rank consistency of each node. Different procedures are executed for performing rank consistency.

RE-VAL. Rank and energy validation (RE-VAL) is executed through DIO messages. After receiving the DIO messages, it validates the information to check whether it is consistent or not.

RE-VER. After validating the rank and energy, this procedure rank and energy verification (RE-VER) is used to verify the RE metrics of each node. The node's rank value is calculated through any one of the objective functions.

SUB-RInA. The procedure called substantiation of rank inconsistency attack (SUB-RInA) identifies the malicious node by the evidences collected through RE-VAL and RE-VER procedures. The malicious node will then be eliminated from the network using any one of the repair mechanisms.

3.4 RPL-REPAIR

Repair mechanisms are important procedures of RPL routing protocols. The RPL-REPAIR procedure works dynamically in the context of the node or link failure. When a root node detects rank inconsistency between two nodes, it uses either local repair mechanism or global repair mechanism.

4 Components of the AROSTEV Framework

The proposed framework AROSTEV is developed to provide secured routing for IoT environment. It detects and mitigates the major attacks such as rank decreased attack (RDA), rank increased attack (RIA), and rank inconsistency attack (RInA) which are similar to selective forwarding, sinkhole, and blackhole attacks. Also, AROSTEV framework enhances the PDR and reduces packet transmission delay. Here, the attacker node progresses with its malicious activities to disrupt the communication. This type of change in the system makes the network vulnerable. So, the data packets do not reach the destination properly. The root node uses AROSTEV framework to handle these types of issues (Table 1).

So, as to recognize the rank inconsistency and to ensure that the hub's rank is steady over the organization, the 6BR checks every node's position, energy, and compares the data with its neighbor hub's data and distinguish the inconsistencies. The 6BR can raise a false alert when the discovery of hub's rank is conflicting. The calculation utilizes the acknowledge of energy as a boundary to discover the attacker hub which has bogus data. The calculation depends on two conditions: (i) checks the revealed estimation of individual hub's position and energy, (ii) contrasts similar detailed worth and its neighbor hubs.

Figure 2 depicts the simulation of AROSTEV framework. The 6BR acts as a gateway between the connected devices or sensors and the Internet. The major role is to connect devices with the Internet. In Fig. 2, the sensor nodes are connected,

Table 1 Algorithm for finding rank inconsistencies

```
Require: N- A list of nodes
for node in N do
if
        Rank (Node) = Rank (Parent)+ MinHopRankIncrease
          Rank (Node) > Rank (Parent)
       Require PDR (Set of nodes)
then
          Select the preferred parent node
end if
end for
for node in N do
     if
         Rank (node) < Rank (parent)
       Rank inconsistency occur
        then
          Require repair mechanisms
    end if
end for
```

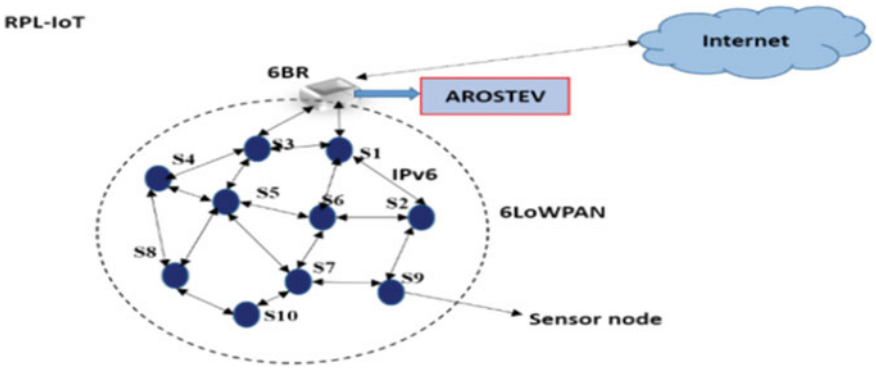


Fig. 2 Simulation on AROSTEV in the 6BR

represented as S1, S2, S3…S10. These nodes sense the data, process, and finally deliver it to the centralized device or gateway; in which, the attacker node can capture the data by using different types of attacks. The AROSTEV framework is deployed on the centralized device (6BR) to identify the attacker nodes' progress and mitigates them.

The AROSTEV framework is based on the attacker nodes' events such as RDA, RIA, and RInA attacks. Similarly, it also has the ability to identify and correct the loop in the network. Consequently, AROSTEV ensures the secure routing in RPL-based IoT.

5 Simulation Results

Here, the experimental setup and the assessment of AROSTEV framework are presented based on the various network parameters which are given below in (Table 2).

Figure 3 shows the process of RIA attack with 13 nodes. The attacker node's ID as well as rank increment process are displayed in the motes output screen.

Figure 4 shows the identification of RIA attack and the loop. The proposed sink node traverses the entire network and identifies the RIA attacker node and loop

Table 2 Parameters for simulation

Simulation parameters	Description
Total number of nodes/hubs	50,100,150,200
Simulation range	1000 m × 1000 m
Node arrangement	Random, grid
Radio medium	UDGM-distance loss
Operating system	Contiki

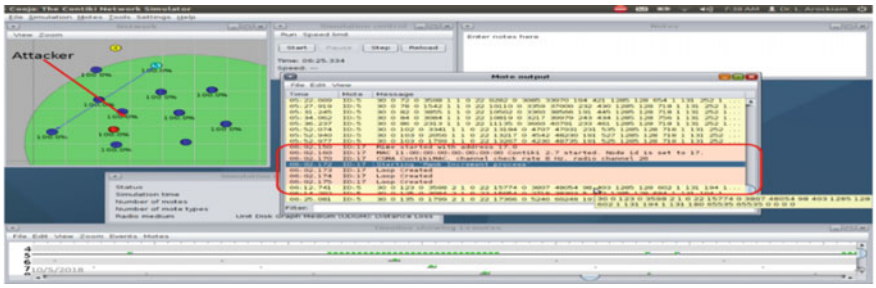


Fig. 3 Process of RIA attack

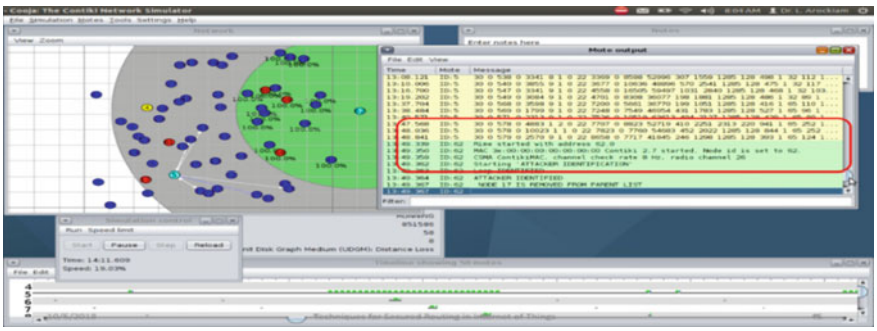


Fig. 4 Identification of RIA attack and loop

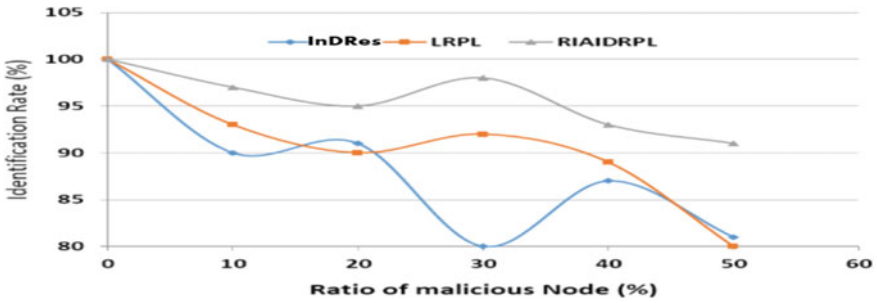


Fig. 5 Identification rate of malicious node

location. Each node has a unique ID in the network. So, the proposed sink node easily identifies the RIA attacker node as well as the loop with the ID.

The RIAIDRPL measured the identification rate of malicious nodes among 100 nodes. This measurement uses the different ratio of malicious nodes, also compared with existing systems such as LRPL and InDRes. Figure 5 shows the RIAIDRPL better performances than the existing systems.

The proposed method measures RIA detection accuracy with various numbers of nodes. The AROSTEV achieved almost 95% of the RIA detection accuracy. This accuracy is evaluated with different ratios (10, 25, 50%) of RIA with 10, 50, 100 number of nodes which is shown in Fig. 6.

The AROSTEV checks the packet delivery rate (PDR) in milliseconds (ms). It achieved almost 90% of packet delivery rate at 2000 ms. This measurement is compared with the existing system LRPL. From this result, it is clear that the AROSTEV gives better performance in PDR ratio as shown in Fig. 7.

The AROSTEV measures the packet transmission delay with various attacker ratios such as 0.2, 0.4, 0.6, 0.8, and 1%. It measures the delay in seconds. Here, the AROSTEV reduces the packet transmission delay compared with the existing system LRPL. Figure 8 shows the evaluationof AROSTEV with existing LRPL by keep on changing the ratio of attacker node.

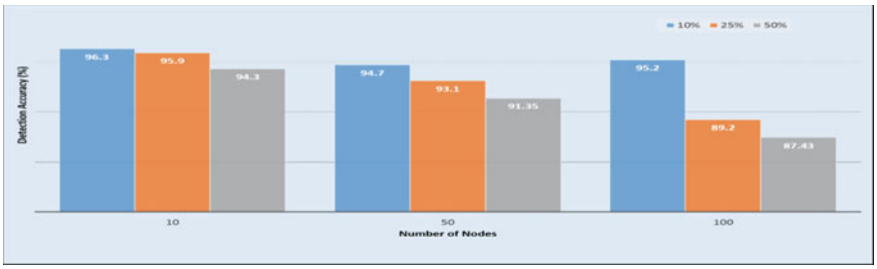


Fig. 6 RIA detection accuracy

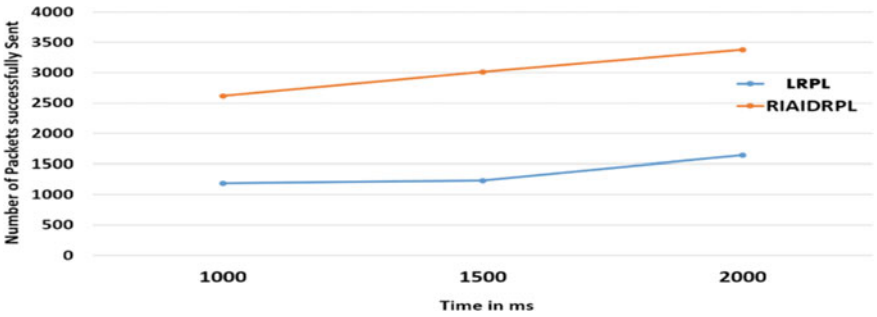


Fig. 7 Packet delivery rate

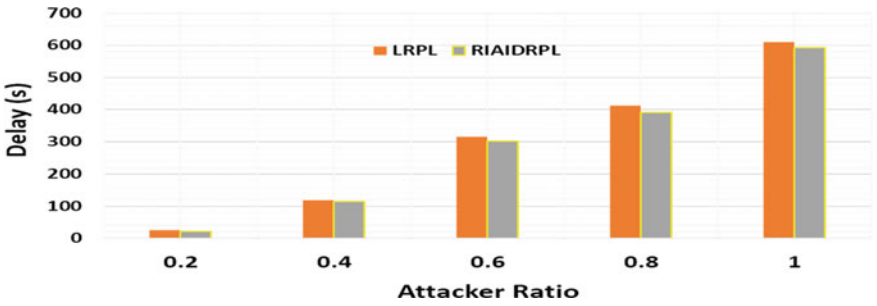


Fig. 8 Packet delay: AROSTEV with LRPL

6 AROSTEV: Deploying in Smart City Application

The IoT applications are becoming widely used in day-to-day life. The proposed framework AROSTEV meets the requirements of IoT applications. This framework is specifically designed for IoT network. The MSP430 is used as a tool for formulating the real-time applications which integrate IPv6 connectivity and 6LoWPAN. It integrates sensor and software development kit (SDK) to test routing attacks. This section employs the AROSTEV framework in the real-world IoT application using MSP430 node. The MSP430 has the capability to communicate with other entities through Wi-Fi module and CC2420 radio.

The nodes are fortified with sensors CC2420 radio, and a battery to bring together data and transmit them to the gateway. The gateway then sends data to the server by using IPv6 interface. The AROSTEV framework is implemented in MSP430 platform and used in the root node for detecting malicious behaviors. Figure 9 depicts the deployment of the AROSTEV framework in the smart city network.

In Fig. 9, various types of sensors are deployed which is used to sense the real-time environment. For instance, noise maps sense the acoustic levels in the city; air quality detects the particulates and dust in the air; structural health monitoring senses the crack detection and propagation; waste management measures the garbage

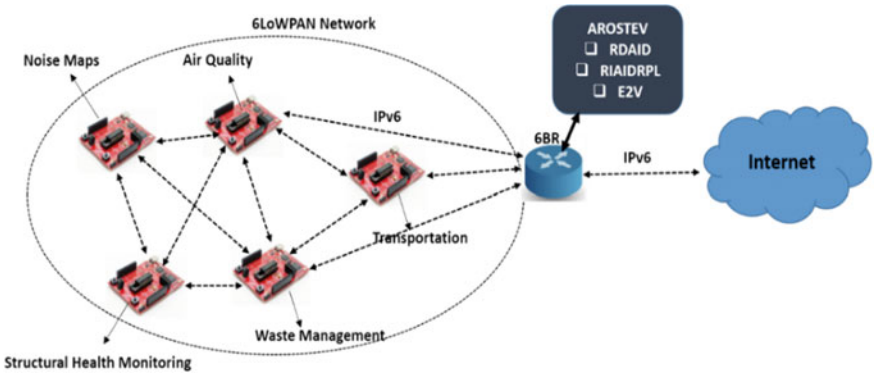


Fig. 9 Deploying AROSTEV in smart city application

levels, and transportation system senses the road condition or traffic. The sensors can communicate and share their sensed data together.

Finally, all the data are reached to the gateway through the intermediate nodes or router. The gateway 6BR acted as a bridge between the connected sensors and the Internet [10]. The AROSTEV framework is deployed in the 6BR node to analyze the behavior of sensor nodes in the 6LoWPAN network. The AROSTEV framework consists of RDAID, RIAIDRPL, and E2V techniques. Therefore, the AROSTEV framework makes the network secured.

7 Conclusion

A framework AROSTEV has been developed to provide secure routing to the RPL-IoT environment. The framework AROSTEV integrates the proposed three techniques, namely RDAID, RIAIDRPL, and E2V. Each technique takes the responsibility to progress its activity against the internal routing attacks. The AROSTEV framework is executed during the RPL DODAG rank inconsistency. AROSTEV enables two types of repair mechanisms based on the level of inconsistency found in the RPL DODAG.

References

1. Shahid R, Wallgren L, Voigt T (2013) SVELTE: real-time intrusion detection in the internet of things. Ad hoc Netw 11(8):2661–2674
2. Christian C, Poplade D, Nogueira M, Santos A (2015) Detection of sinkhole attacks for supporting secure routing on 6lowpan for internet of things. In: Integrated network management (IM), IFIP/IEEE international symposium on IEEE (2015)

3. Surendar M, Uma Makeswari A (2019) InDReS: an intrusion detection and response system for internet of things with 6LoWPAN. In: Wireless communications, signal processing and networking (WiSPNET), international conference on IEEE, pp 1903–1908
4. Stephen R, Arockiam L (2018) RDAID: rank decreased attack identification algorithm for internet of things. Int J Scientif Res Comput Sci Appl Manage Stud (IJSRCSAMS) 7(3). ISSN: 2319–1953
5. Stephen R, Arockiam L (2018) RIAIDRPL: rank increased attack identification algorithm for avoiding loop in the RPL DODAG. Int J Pure Appl Math (IJPAM) 119(16):1203–1210. ISSN: 1314–3395
6. Stephen R, Arockiam L (2018) E2V: techniques for detecting and mitigating rank inconsistency attack (RInA) in RPL based internet of things. J Phys: Conf Ser (JPCS) 1142:75–86. IOP publishing, ISBN: 978–93–5311–592–0
7. Amit D, Buttyan L (2011) VeRA-version number and rank authentication in RPL. In: 8th international conference on mobile adhoc and sensor systems (MASS), IEEE, pp 709–714
8. Jianlin G, Han C, Orlik P, Zhang J, Ishibashi K (2012) Loop-free routing in low-power and lossy networks. In: SENSORCOMM 2012. The sixth international conference on sensor technologies and applications, pp 59–66
9. Anuj S, Mayzaud A, Badonnel R, Chrisment I, Schönwälder J (2014) Addressing DODAG inconsistency attacks in RPL networks. In: Global information infrastructure and networking symposium (GIIS), pp 1–8
10. Dar AR, Ravindran D (2019) Fog Computing: An efficient platform for the cloud resource management. J Emerg Technol Innov Res (JETIR) 6(2):7–12 ISSN: 2349–5162
11. Heiner P, Landsmann M, Ugus O, Schmidt TC, Wahlisch M (2013) TRAIL: Topology authentication in RPL

Early Prediction of Hydrocephalus Using Data Mining Techniques

S. V. Rajiga and M. Gunasekaran

Abstract Hydrocephalus is a condition which is characterized by head enlargement in infants due to enlargement of brain ventricles. An excess of fluid secretion and collection of fluid within the brain cavities are treated as Hydrocephalus. The extra fluid puts stress on the brain and can damage the brain. Hence, the increase in the fluid level in the brain's cavities may increase intracranial pressure and lead to brain damage. It is most usual in infants on children and rarely in the adult age group. Children often have a full life span if hydrocephalus is early detected and treated. This paper presents various data mining techniques and used to find out the disease in an early manner. Magnetic resonance imaging is one of the detection tools which is used to predict the disease properly. This approach includes the basic four data mining processes, namely preprocessing, segmentation, feature extraction, and classification as stage-by-stage manner using MRI dataset. Along with this process, the tree augmented Naïve Bayes nearest neighbor (TANNN) algorithm is also implemented to improve the accuracy in detecting the disease and also gave the best detection rate. The TANNN algorithm may provide the best results in diplomatic, uniqueness, perfection, and overall running time. The first stage in the data mining technique is preprocessing, which converts the original data into a useful format. The second stage is a key technique, and it groups the original data into possible divisions according to its category. The third stage is feature extraction, which is used to extract the needed data from the source. The fourth stage is the classification that appoints data in a collection to destination categories or data groups. This paper also concentrates on image mining, which includes experiments in image elements such as texture, shape, and size. Image classification is an important task in the field of medicine and technology. This helps the radiologist in the process of diagnosing hydrocephalus.

Keywords Hydrocephalus · Data mining · Preprocessing · Segmentation · Feature extraction · Classification

S. V. Rajiga (✉)
Research Department of Computer Science, Government Arts College (Autonomous), Salem 636007, India

M. Gunasekaran
Research Department of Computer Science, Government Arts College, Dharmapuri, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_26

1 Introduction

Early detection of hydrocephalus is most important in clinical practice. "Hydro" means water, and "Cephalus" means head. Hence, hydrocephalus means incremented cerebrospinal fluid (CSF) in the brain. The common treatment in hydrocephalus is the surgical placement of ventriculo peretonial shunt (VP shunt).

It is a long flexible tube used to expel excess cerebrospinal fluid from the brain into the peritoneal cavity. One end of the tube is put down in one of the brain's ventricles, the other end into the peritoneal cavity [1].

Hydrocephalus is a chronic condition. If properly monitored means, it can be controlled but not completely cured. Classification of hydrocephalus is based on the abstraction of CSF flow. That is communicating and non-communicating hydrocephalus and abstractive non-communicating hydrocephalus. Causes of hydrocephalus in children include spina bifida, sludge I and II deficiencies, arachnoid cysts, and generally hydrocephalus stenosis.

It is easy to diagnose hydrocephalus, but the aim should be to find out the exact lesion. This article aims to analyze the uses of data mining techniques in detecting hydrocephalus for achieving high accuracy. The four data mining methods used are preprocessing, segmentation, feature extraction, and classification as shown in Fig. 1. The tree augment Naive Bayes classification technique is being implemented to diagnose and classify hydrocephalus in children [2].

In the hydrocephalus prognosis and detection step, the hydrocephalus area is calculated using the binarization method, which converts the dark pixel into dark and white pixel whiter. Input coding system, each transformation coefficient is compared to a threshold; if it is less than the threshold value, it is considered zero or otherwise one. In the rational phase, the hydrocephalus fraction is calculated using the binarization method. This is also identified and analyzed.

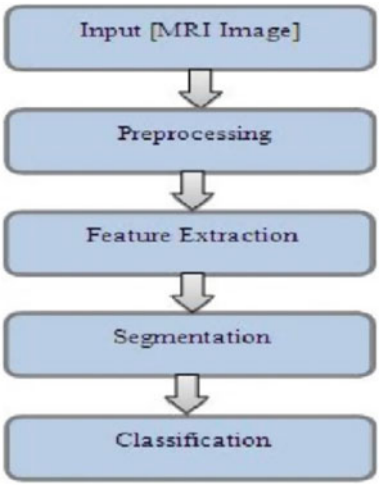


Fig. 1 Block diagram depicts the data mining techniques in hydrocephalus detection

2 Methodology

According to the need for diagnosis, four basic data mining techniques with one algorithm help increase the accuracy in diagnosing the disease. The algorithm makes several classifications based on the learning theory to determine the probability that one algorithm belongs to certain [3]. At the diagnostic and classification stage, the tumor area is calculated using the darker pixel and white pixel brightening method. The feature is the cluster's extraction, which shows the predicted disease and is delivered to the extracted cluster gate process. It uses a binary mask for the whole image. It is an image with only two values, black or white (0 or 1).

The exact methodology as shown in Fig. 2 explains four basic data mining methods, one algorithm, and its usages in detecting dangerous diseases like hydrocephalus.

2.1 Image Preprocessing

Image preprocessing is a technique for improving image data before computational processing. Preprocessing aims to improve the image data, which suppresses unnecessary distortion or further enhances certain image features that are important for processing [4]. The following steps are shown in Fig. 3.

The original image MRI scan is taken for the preprocessing process. Finally, the image without noise is taken for the next process.

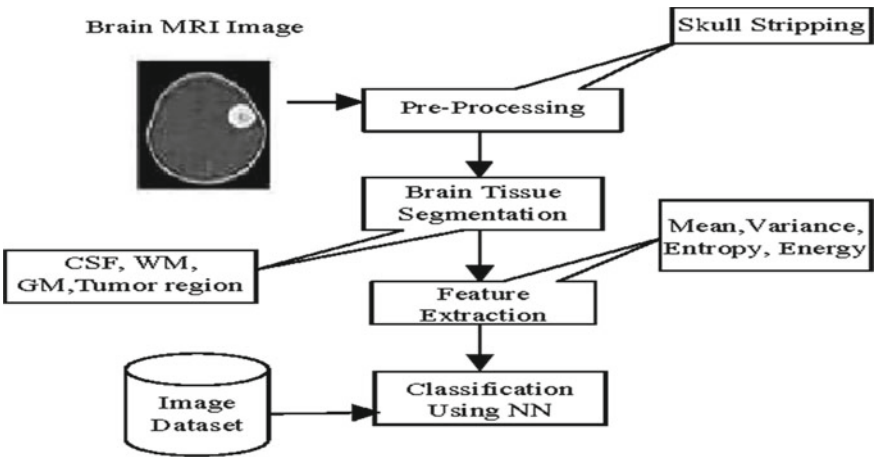


Fig. 2 Block diagram of the proposed hydrocephalus system

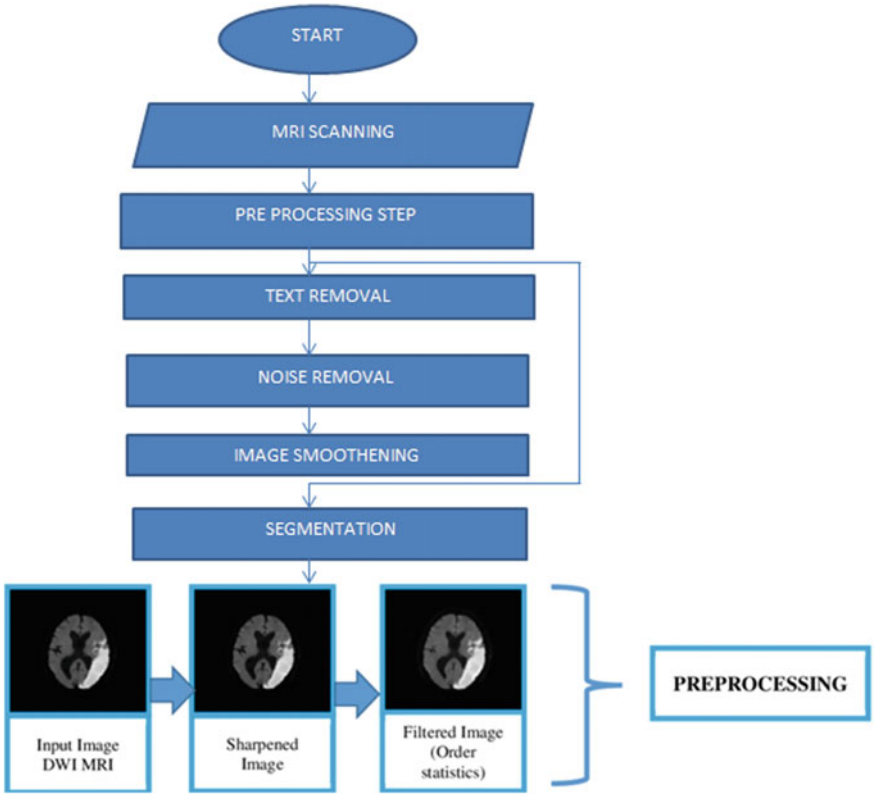


Fig. 3 Block diagram of preprocessing steps

2.2 Segmentation

Segmentation refers to the process of separating information according to its category. It is the process of splitting a digital picture into several parts. It is commonly used in images to find objects and boundaries. The image section is the result of a collection of sections covering the whole image collectively as in Fig. 4.

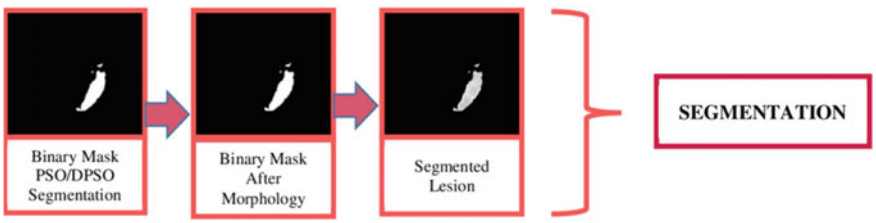


Fig. 4 Block diagram of segmentation steps

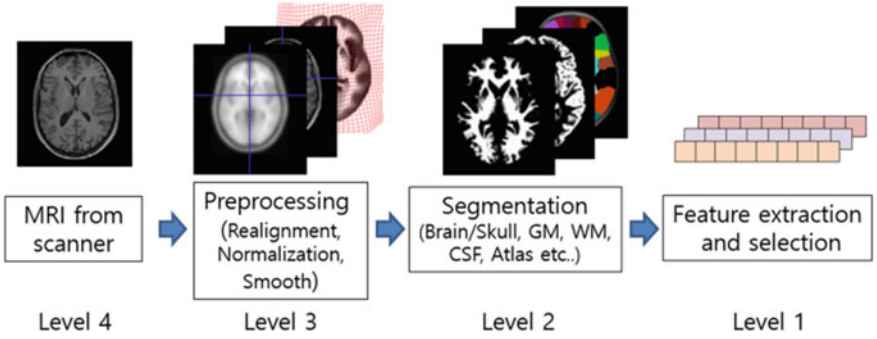


Fig. 5 Block diagram of feature extraction

The image segment's goal is to divide an image into a set of homogeneous and non-foldable areas that have meaningful properties such as intensity, depth, color, or texture [5].

2.3 Feature Extraction

Feature extraction reduces the size, thereby reducing the initial set of raw data into manageable groups for processing as hown below in Fig. 5.

The purpose of the extraction of features is to reduce the original data by measuring certain features or characteristics that distinguish one input format [6]. The following features of the system are extracted as shape, intensity, and texture features.

2.4 Classification

Classification is a data mining function that targets the categories or classes of objects in a package. The objective is to accurately predict the target class in the data for each case. In image classification, image classes should clearly define the purpose and characteristics of the image [7]. The purpose of the classification is to identify and portray a unique entity.

In this classification module as shown above in Fig. 6, it is analyzed that whether the disease is because of picture achieving and communicating system (PACS), lacunar stroke (LACS), total anterior circulation stroke (TACS), etc.

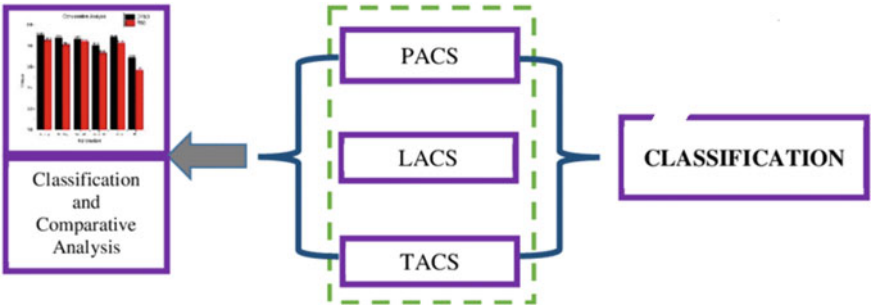
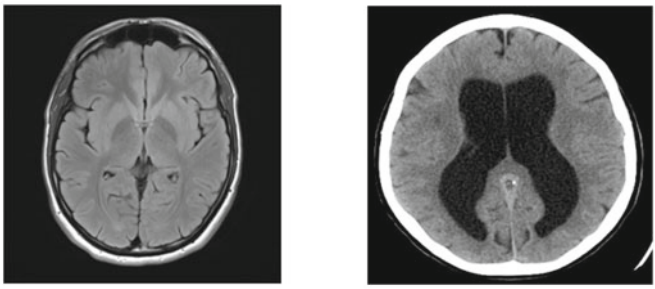


Fig. 6 Block diagram of classification

3 Discussions and Experimental Results

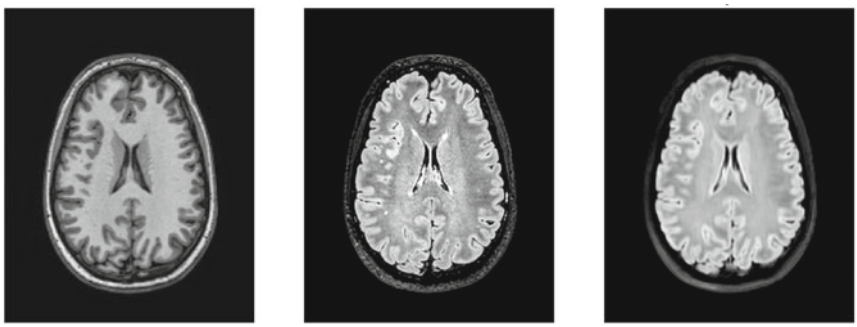
In the above section, the results obtained using MRI scan are presented. All the four basic data mining techniques are applied one by one on MRI scan image. Finally, the TANNN algorithm is implemented to improve its accuracy in detecting the disease (Fig. 7).

The most popular medical imaging diagnosis in the medical field is called MRI. MRI is a method of medical imaging used in radiology to create an image of the body's anatomy and physiological processes [8]. To create images of organs in the body, MRI scanners use strong magnetic fields, magnetic field gradients, and radio waves.



(a)Normal Brain Image (b)Hydrocephalus Brain Image

Fig. 7 The figure depicts the structure of the normal brain and hydrocephalus



(a)Brain Image (b)After Noise Removal (c)After Smoothening

Fig. 8 Two-steps noise removal and smoothening represented

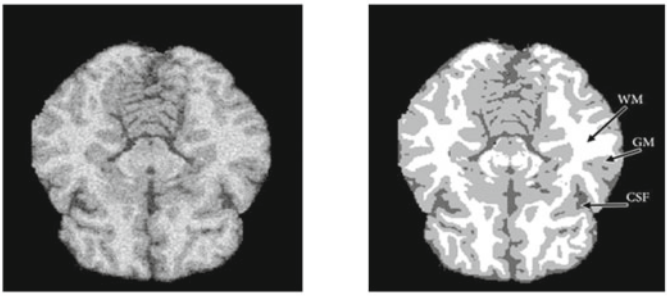
3.1 Preprocessing

The first module preprocessing is applied to MRI scan images, and the result is presented. Each brain image represents each step of preprocessing. Here, two steps are noise removal, and smoothening is applied as shown in Fig. 8.

3.2 Segmentation

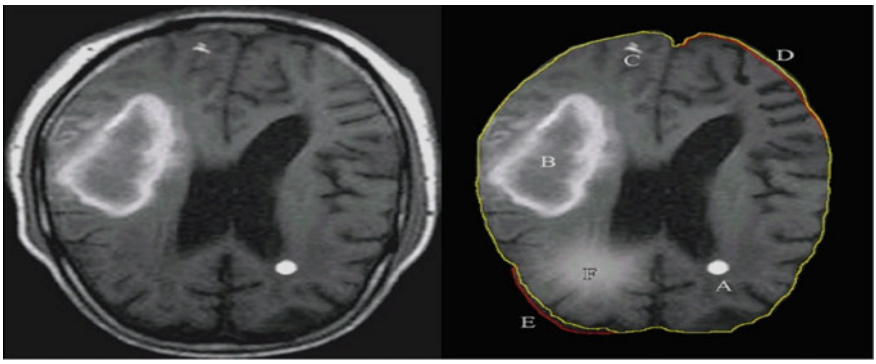
The second module segmentation is applied to MRI scans. The result presented is divided the brain image into its segments of particles as shown in Fig. 9.

The first image was detected. It is divided into three categories as white matter, gray matter, and cerebrospinal fluid [9]. The final CSF level will only diagnose hydrocephalus.



(a)Brain Image (b)After Segmentation

Fig. 9 Segmentation process



(a)Hydrocephalus Brain Image (b)Extracted Portions Identified

Fig.10 Feature extraction process

3.3 Feature Extraction

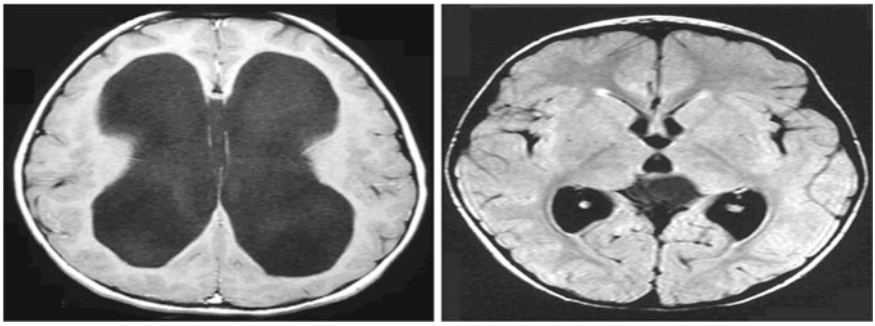
The third module feature extraction extracts the needed portion from the abnormal brain image and identifies the problematic cells as shown below in Fig. 10.

In the above figure, the first one depicts the hydrocephalus brain image. The second one depicts the identified portions indicated with labels A, B, C, D, E, and F. This is how the portions are extracted with the help of the feature extraction process in the data mining technique.

3.4 Classification

The fourth module and the important module are used to identify the exact disease in the brain and the exact portion affected in the brain (Fig. 11). The exact disease diagnose is happened here. The classification system consists of a database, and pre-defined shapes compared to the object found to be classified as the correct type [10]. Image classification is an important and challenging task in various applications, especially in the medical field.

In this section, only tree augmented Naïve Bayes nearest algorithm is used to bring the multi-classification concept. It is very easy and is used to find the probability of an image belonging to a particular class. Its purpose is to categorize the image with minimal time difficulty. After this process, the neighboring neighbor is used to classify the brain's normal and abnormal part.



(a)Hydrocephalus Brain (b)CS Fluid excess level indication

Fig. 11 Classification process

4 Conclusion

In this paper, it is analyzed how various data mining concepts are used to detect hydrocephalus in the step-by-step manner. It is studied that how the TANNN algorithm is used in the early detection of hydrocephalus under the classification concept. This paper provides a brief overview of image mining, which is an extension of data mining in image processing. This study preprocesses an important and fundamental role in predicting the right disease in the right manner using MRI of the brain; it also examines how brain tissue such as white matter, gray matter, and cerebrospinal fluid can be separated. The purpose of this study was to investigate whether features extracted from MRI in predicting the disease. The final process of classification analyzes the place of the disease and the occurrence of the disease. Understandably, these basic methods of data mining are very much useful in predicting dangerous diseases like hydrocephalus. The method mainly includes preprocessing, segmentation, feature extraction, classification with TANNN, and image mining.

References

1. Manoj TH, Gunasekaran M, Jaisingh W (2019) A brainnet classification technique based on deep convolutional neural network for detection of brain tumor in FLAIR MRI images. Int J Eng Adv Technol 9(1):3264–3269, Blue Eyes Intelligence Engineering & Sciences Publication
2. Sathish Kumar R, Nagarajan B, Karthikamani R, Gunasekaran M (2017) Region-based object extraction using adaptive neuro-fuzzy inference system combined with support vector machine. Asian J Res Social Sci Humanities 7(2):412–427. Asian Research Consortium
3. Ali EM, Seddik AF, Haggag MH (2016) Real brain tumors datasets classification using TANNN. Int J Comput Appl 146
4. Lehmann TM (2005) Automatic categorization of medical images for content-based retrieval and data mining. www.elsevier.com

5. Swati ZK, Zhao Q, Kabir M, Ali F, Ali Z, Ahmed S, Lu J (2019) Content-based brain tumor retrieval for MR images using transfer learning. IEEE Translation and Content Mining are Permitted for Academic Research 7
6. Pushpa Ravi VPG, Palani S (2012) A novel approach for feature extraction and selection on MRI images for brain tumor classification. J Comput Sci Technol
7. Jaisingh W, Duraisamy M (2018) A novel algorithm for detection of lesions in digitized mammograms using fuzzy C—means bootstrap with convolutional neural networks. Int J Pure Appl Mathem 2585–2590
8. Bahadure NB, Ray AK (2017) Image analysis For MRI based brain tumor detection and feature extraction using biological inspired BWT and SVM. Int J Biomed Imaging
9. Abirami N, Karthik S, Kanimozhi M (2015) Automated brain tumor detection and identification using medical imaging. Int J Res Comput Appl Robot 3
10. Despotovic I, Goossens B (2015).MRI segmentation of the human brain: challenges, methods, and applications. J Comput Mathe Methods medicine 2015

An Insight View on Denial of Service Attacks in Vehicular Ad Hoc Networks

Naramalli Jaya Krishna and Narayanan Prasanth

Abstract Vehicular ad hoc network (VANET) is a modern communication design that allows the interpretation between vehicles to move at high speeds on the roads. It has started the opportunity to design various new administrations like traffic administration, distribution of emergency data to avoid dangerous conditions, traffic engineering, and other user records. VANETs are a close relation of MANETs but with recognizing features like action at high speeds, processing power, moving controlled, sufficient storage, variable node consistency, and challenging communication position with precise link existence. Various security issues are usually identified in vehicular ad hoc networks, like authentication and authorization issues, and it gets problems from frequent attacks that survive for MANET. As VANETs are extensively used in various critical applications, much investigation has been done to determine effective solutions against these DoS attacks that can work for multiple routing protocols. This paper presents a brief survey on DoS attack on network availability and its extremity stage in the VANET atmosphere. Also, the paper discusses on various existing techniques to detect and prevent the DoS attack. A detailed comparative statement is made between these techniques to understand its effectiveness in VANET.

Keywords Vehicular ad hoc networks · Distributed denial of service attack · Security · VANET routing

1 Introduction

Vehicular ad hoc network (VANET) is a kind of network where the vehicle nodes can speak in a mode of crossover with each other on the road [1]. VANET packages

N. Jaya Krishna (✉) · N. Prasanth
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

N. Prasanth
e-mail: n.prasanth@vit.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_27

are widely marked in security and non-security programs. Safety packs are essential, as they relate to users and their lives without delay. These packages provide records related to drivers' warnings, including block notification on a particular street. But merely, VANET challenges to protect human lifestyles, even when these humans move on the roads. Unprotected packages are the convenience of drivers and passengers, enhancing the visitor's device. Examples of these packages include travel map, parking availability, and weather data.

Generally, both types of programs are intended to provide accurate statistics to clients/drivers on the road. However, for protection applications, the information that is no longer effective must be accurate but is also safely transferred from source to vacation location. Therefore, security is a significant problem as small interruptions, along with intermittent disconnection, can create problems for customers. It was particularly necessary if the constant presence of data was communicated between both communication peoples. To this, network availability was an essential element. As identified in [2], accessibility is a significant security element. It was described as if any node wants to gain acceptance of the alternate node within the network or access to the structure, and the center should be helpful to the communication nodes. Inaccessibility can provide to any failure or any harm, such as a DoS attack.

Security issues in VANET are critical capabilities in today's scenario due to the rapid boom and growth of the VANET app. The most extreme attacks in VANET are a DoS attack. Because it attacks network availability causing an effect threatening, the presence of motorists must be identified to prevent such attacks as quickly as possible since the attacker's primary goal is to disable the communication channel or confuse car offers available from unique customers. The attack makes the device useless, and the device's uselessness in vehicle networks in real-time, even for a diminutive time, is risky.

VANET implementation is involved in the development of network protocols that can ensure safe and effective V2V and V2I interactions. VANET is sensitive to a variety of factors from one visitor to another, traffic jams, and frequency deviations from time to time on the corresponding routes. Thus, VANET network protocols must be scalable to maintain large networks. VANET has excellent features that require new problems in community production and maintenance when paired with regular MANETs. Unlike traditional systems, where nodes are idle or pass one after another at low speeds, the nodes in VANET transfer are at very excessive speeds, resulting in the segmentation of the community and rapid changes in the network structure. Moreover, the development of the vehicle's contract is subject to the control of the driving force and conversation with nearby cars.

The VANET timeline is designed to provide a systematic way of meditating with tools and equipment close to EE. The USA is a piece of important equipment for near and permanent opportunity problems in Fig. 1.

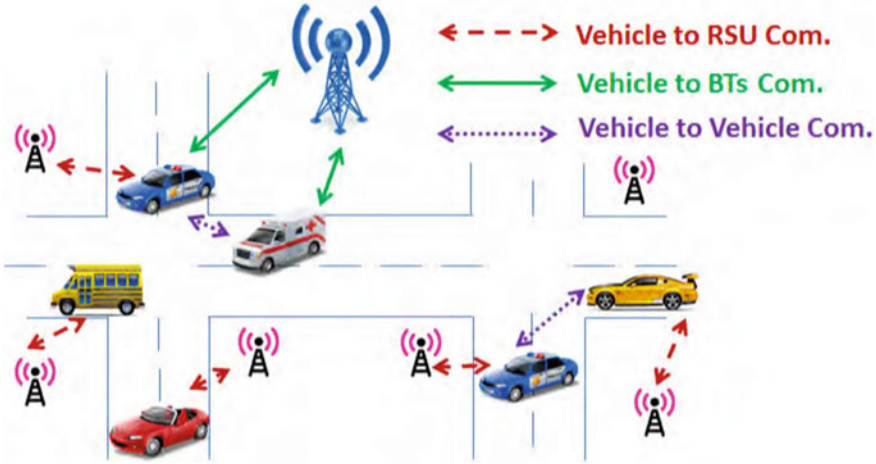


Fig. 1 Traditional VANET architecture [2]

1.1 Vehicle-to-Vehicle (V2V) Communication

If the communication infrastructure is missing, it allows cars to communicate directly between new vehicles. Only onboard units (OBUs) are essential. OBUs are a vehicle-mounted project for data exchange that connects to adjacent spare cars, from V2V or between vehicles or between permanent devices (known as roadside units (RSUs)), i.e., powered vehicle-to-infrastructure (V2I).

1.2 Vehicle-to-Infrastructure (V2I) Communication

It allows the viewer to communicate only via road units. The RSU is a device that is mounted on specific roads and serves as an information center. Roadside units (RSUs) are responsible for offering safety and insecurity tasks to onboard nits (OBUs) and roadside units (RSUs) only.

1.3 Hybrid Architecture (V2V–V2I)

It is a mixture of two procedures, especially from V2I and V2V. The vehicle can match between road infrastructures and vice versa in this approach, with single hop or multiple hop moves. The Federal Communications Commission (FCC) has approved a wireless protocol called as a dedicated short-range communication (DSRC). The

term DSRC is known as a 5.8 GHz frequency band with a full bandwidth of 75 MHz for intelligent transport system (ITS) [3].

The organization of the paper is as follows: Sect. 2 provides the related works on VANET security. Section 3 discusses on DoS attacks in VANET environment, and Sect. 4 discusses on various solutions for DoS attacks and also compares it. Section 5 concludes the paper.

2 Literature Survey

VANET works in a challenging communication environment, which has so far limited sensitive generation deployment. VANETs are particularly vulnerable to hidden node difficulties. Moreover, they must deal with limited spectral bandwidth and an imaginative variable channel affected by the continuous cellular interruption and sources of interference, and infrastructure-based networks have a significant advantage over ad hoc networks. In such a situation getting access to factors, access allows for the scheduling and distribution of network resources to the maximum extent in a simple way. Someone needs to be posted responsible for the needy. To get several infrastructure benefits in the first place without infrastructure, researchers studied clustering in VANET, as some clustering algorithms distribute the entire device. There are many works done in the past to protect vehicle ad networks. Some diagrams are described below,

Bariah et al. [4] have investigated the newly implemented safety arrangements for VANET. As their investigation includes many threats (disclaim, puncture, spam, re-modeling, jamming, black hole attack, DoS, DDoS attack, etc.) and provided solutions, their research showed that threats could be classified according to the V2V and V2I scenario. Also, authors were compared various simulation tools like NS-2, NS-3, NCTUns, and Qualnet, GrooveNet, TraNs.

Singh et al. [5] presented various algorithms to find a DoS attack on VANET; it was called enhanced attacked packet detection (EAPD). In this algorithm, the authors used threshold values and intervals. Their simulation results showed that improved network performance and not provided false notifications.

Saranya et al. [6] studied on the wormhole attack, DDoS attack especially. They compared between various modern attack prevention implementations. One of the studied provided a technique called as FireCol method could reduce the hardness of network damage like a traffic matrix, and it could be used to show peer-to-peer-based area aggregates to site visitors sufficiently. Also, they used bloom filters to preserve routing data. Their survey also included a simulation evaluation of suggested methods.

Mejri et al. [7] investigated the importance of DoS attack in VANET and provided a mechanism based on game theory algorithm. In their works, a game theory algorithm that was employed to the protection of VANETs could be an excellent method to perform. The proposed algorithm watches the board disadvantages of each vehicle found within the evolving network and tries to enhance its performance.

Kumaresan et al. [8] implemented a group key authentication method for authentication to support cluster communication in VANET. The keys were produced with w.r.t. today's openings, and information was confirmed and verified to obtain sure the events increase. Interruption could be identified based on time and proximity updates. The simulation results confirmed that could be operated under a state of perception and could enhance overall network performance using PDR resources.

Jamil et al. [9] provided an overview of the communication exchange in VANET as their research discussed network coding methodology for data processing. The process covers indications for network coding, configuration, and modeling mobile environment modeling, object coding, and real-time recording across networks MAC layer present. The authors also considered some security problems described to rule implementation and helpful guidance.

Faghihniya et al. [10] have developed technology for convenient routing through VANET using the AODV protocol which is a reactive routing protocol. Also, it is a topology-based routing protocol adapted to be used in a multi-hop environment. The recommended protocol showed RREQ quantities in the network. The use of broadcast packets within a quick AODV route detection environment is very vulnerable against DoS and DDoS flood attacks.

Ibrahim et al. [11] have proposed a scheme to alleviate the lexical message communication of V2V and V2I. Many researchers provided security provided message exchange mechanisms in vehicles. The proposed system presented unique strategies, i.e., the service-oriented architecture (SOA) infrastructure. It could be a new choice to provide all security requirements. Their experimental showed that their proposed infrastructure was better than the previous provided models.

Dong et al. [12] states the DDoS attack against controller is one of the safety threat of software-defined networking. In the scenario of DDoS attacking, the attackers could make substantial new low traffic flows to the network to disturb the system. Here, the intruder detects the notion of monitoring and employs serialization ability test to compute false/real frequent errors. Their results showed the overall performance in inaccurate terms and can advance multiple transmissions.

Rehman et al. [13] presented rule-based mechanism to find out the DoS attack for that they deployed a way to support Internet Protocol version (IPv6). The console uses IP tables and temporary logging to analyze upcoming site visitors. Database usage has been tested on all IP addresses, and if any of the IP addresses is abnormal, it is very far from the network. Their results showed overall performance in terms of accuracy and precision.

3 Denial of Service Attack (DoS)

The attacker frequently disturbs the communication network in the VANET environment by jamming or creates nodes issues to approach the system. The primary goal was to limit real nodes from obtaining network services or using network assets due to this node, unable to receive or send essential information in the network. Finally,

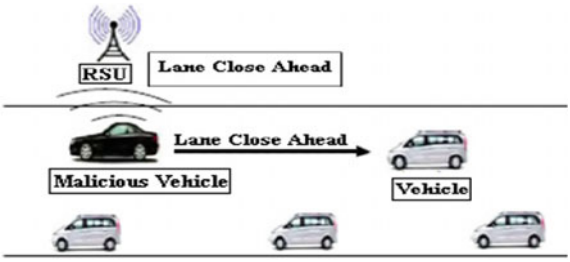


Fig. 2 Example of DoS attack [6]

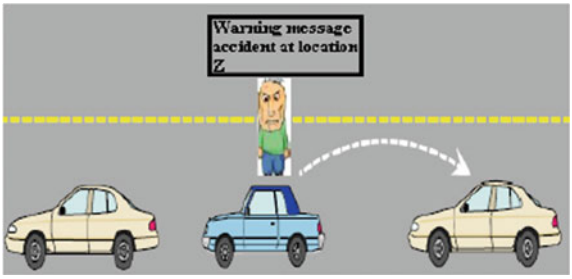


Fig. 3 V2V communication at DoS attack [14]

the network is not available for a long time for customers. DoS cannot be permitted to appear on VANET because having relevant information needs to get to your vacation place safely and on time. There are three ways the accused can also get DoS attacks: talk channel interference, overhead network assistance, and packet loss. An example of a DoS attack is shown below (Figs. 2 and 3).

Case 1: V2V Communication suffers by DoS attack:

According to the third figure, the attacker sends a "fate of prevention in Z" alert. The statistics behind the attacking node will receive these statistics. The attacking node will send this identical message again and again to the suffering node, making it busier and denying you access to the network [14].

Case 2: DoS attack in Vehicle-to-infrastructure (V2I) communication.

At this level of attack, the attacker launches attacks on the road unit. When RSU is busy reviewing messages, no other node that wants to change facts with RSU will receive any reply. Therefore, network services are not available for this node (Fig. 4).

Another worse case in a DoS attack is called a distributed denial of service (DDoS). Here, there are more than one attacker. Therefore, this attack is widespread and difficult to protect. The attacker uses different locations/geographical locations to launch this attack [15]. Two cases of this assault are:

Case 1: Here, attackers attack from many locations with many times. Therefore, the shape of the sent message and the time interval may differ from one node to another (Fig. 5).

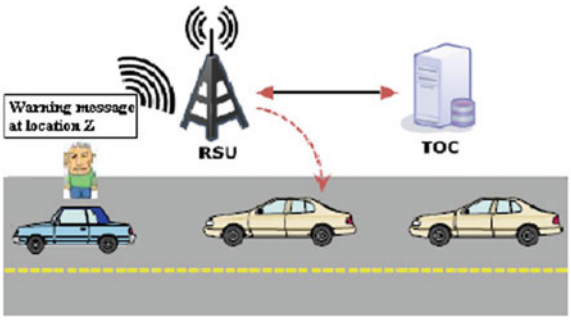


Fig. 4 DoS attack in V2I communication [7]

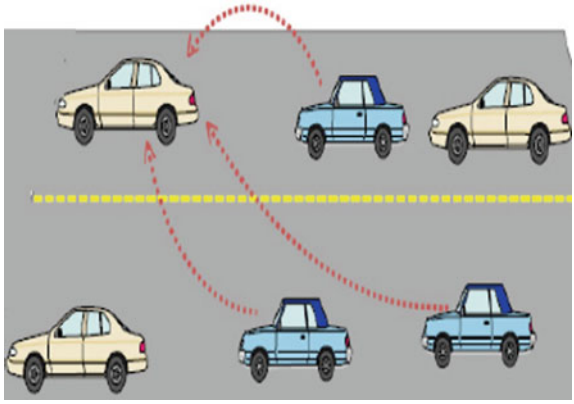


Fig. 5 DDoS attack in V2V communication [8]

Case 2: In Case 2, the attacker intimates the VANET (RSU) infrastructure. In Fig. 6, it appears that three attackers are posing a security risk to the VANET infrastructure with different locations. When the node tries to speak to RSU, it is seen to be overloaded.

4 Various Solutions for DoS Attack in VANET

Nowadays, there are various solutions suggested to provide security from DoS attack in previous.

4.1 Quick Response Tables (QRT)

The scheme changed to a proposal to make VANET comfortable with DoS, using quick response tables (QRTs) to check frequent updates in routing data and benefit

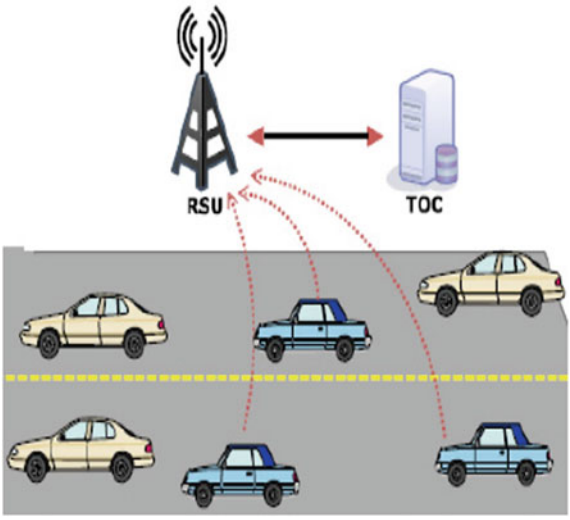


Fig. 6 DDoS in V2I communication [9]

from the reference table. All routing data is recorded according to activities carried out over the network. If there is a decrease in the beam in one cycle and its purpose is unknown, its reduction depends mechanically. After reaching the threshold rate, the present route is isolated from society, if the node drops the beam, without good reason. During the route update section, the use of recognized QRTs, paths, and nodes is ignored and cannot be considered for routing reasons.

4.2 ReCAPTCHA Controller Mechanism

In this mechanism, the legitimate users are restrained from obtaining useful suggestions or devices on the network. The primary objective of the attack is to create the requested destination node destroy or to spread the information across the network affecting accessibility. The proposed reCAPTCHA control mechanism prevents automatic attacks in addition to zombie robots. The reCAPTCHA driver is used to check and limit automated DDoS attacks as possible. To enforce this technique, the entire standardized reporting principle is used to investigate deviation in entropy user demand. Entropy and frequency are the measures used to determine the vulnerability of an attack. A random-copy reCAPTCHA program is used as a safeguard for attackers who rely on giant robots.

4.3 Spoofed IP Addresses (SIPA)

This technique takes advantage of the fact that two nodes will talk for a short time. In the case of the malignant node and the victim, it may be for an unlimited period. Therefore, the statement may extend abnormally long. The author provides the distinction between the attacker nodes and the normal node in the hypothesis of variance in its communication period with the threshold time. Since a DDoS attack almost jumps the offering into exhaustion by sending an abnormally huge group of packets, why not the attacker reflects by testing the increase of packets sent over a node or a skilled blow. In our proposed scheme, we are thinking of making better use of battery resistance assistance, since the nodes in this example are cars. Therefore, we can use the fast and challenging OBU guidelines that can be executed online with the vehicle always to locate the attacker. Consider a random car in VANET, which receives request messages from an IP pair (unique units) at the same time. In a set of rules designed, OBU receives these order messages sent by two sellers and records associated with these senders, for example. Their IP addresses and the packets that they send will be sent to the counter unit.

4.4 Prediction-Based Authentication (PBA)

It is an effective broadcast authentication scheme to protect DoS based primarily on the account due to signature dumping and supporting packet loss in a lost situation. In addition to minimizing signature verification delays, it improves authentication efficiency with Merkle hash tree (MHT) for quick verification of messages, while the targeted guide and TESLA are waiting to protect from abuse adequately. Memory-based DoS, hidden via storage, local puzzle key. To send a poster safely, the sender creates non-public serial keys, waits for the role of vehicles, using web page prediction, and creates a Merkle hash tree to create personal and public keys whose use offers signs and symptoms for the message. If the built-in MAC has identical MAC memories, then the receiver authenticates the guide immediately. To tackle resistance to ray loss, PBA uses the premium TESLA system.

4.5 Ant Colony Optimization

The ant colony optimization algorithm has been suggested to present a realistic, streamlined, and practical method for the convenience of guidance technology by separating a malicious attacker on the way to the vacation spot. Defense against the black hole attack, a form of DoS attack, is competing for an ACO-based directive that is set confidently and at the price of the evil pheromone knot. The inexpensive

node to believe and pheromones are identified as dangerous, and any way to adapt through the node is removed.

4.6 IP Check

A robust and distributed defense technique against DoS attack using IP addressing has been proposed with the attendee in the periodic message sent, which suggests DoS detection while facing related IP processing. The IP address maintained to decrease memory requirement and decreases the detection time that used bloom filter, new information structure.

4.7 FastAuth and SelAuth

This technique proposed two green transmission authentication strategies as a counter-flow to signatures. Fast authentication (FastAuth) secures unpaired beacon messages by predicting destination using (HST) Huffman hash trees. These are used to create a unique signature-based mechanism where the testing time is 50 times more active and generation time 20 times faster than elliptic curve digital signature. Selective authentication algorithm preserves messages from various transfers using nearby data. In wireless loss environments, increased communications and public accounts arise as a final result of packet loss management.

5 Comparison Between Various Solutions for DoS in VANET

The relative comparisons of different solutions for DoS in VANET are discussed in Table 1.

6 Conclusion

Security is the primary aspect of regular road users. The security specifications can be adequately maintained by many protection applications, such as traffic reports generation and accident notification release. The use of VANET can provide such security requirements. Essential messages must be sent from one node to another on the VANET in a reliable and timely manner. Secure communication and network availability must be guaranteed on VANET; however, in most of the cases, the vehicles

Table 1 Comparative study of VANET methodologies

Authors with year	Methodology	Achievement	Pros	Cons
Ying Gao et al. (2019)	A distributed DDoS network intrusion detection system (NIDS) based on big data technology	It gets 0.01% lower FAR rate with the help of ensemble methods use, which is limited in the accuracy	Got the appropriate effects on accuracy and false positive rate (FPR) in detecting DDoS attacks	Unable to implement the real environment for the proposed NIDS. System performance decreases
Charu Guleria et al. (2018)	Greedy approach	Detected and mitigated the DDoS attack	Better throughput, less packet loss, and less routing overhead	In a low-density environment, the most straightforward signatures that lead to a DoS account are used
Paramjit Singh et al. (2017)	Quick response tables (QRTs)	Frequent updates to routing statistics were analyzed and used a reference table	No computational overhead and memory-based burden	Not support for non-reputation, multi-hop conversations, and packet loss no longer help
Poongodi et al. (2016)	reCAPTCHA controller mechanism (reCCM)	It avoids the automated attacks in the same way like botnet zombies	It excluded most of the automated DDoS attacks	Security in low level with DoS attacker
Munazza et al. (2016)	Spoofed IP addresses (SIPA)	SIPA has been suggested to them to distinguish between an attacker and a normal node over the idea of variance in their contact time with threshold time	Protects against IP addresses spoofing and maintains bandwidth	It only detects DoS
Chen Lyu et al. (2015)	Prediction-based authentication (PBA)	Fast verification of beacon messages	It basically resists computing based on memories and is also able to withstand packet loss in a loss environment within the lost environment	In a lost environment, the primary signature of computational costs needs to be addressed

(continued)

Table 1 (continued)

Authors with year	Methodology	Achievement	Pros	Cons
Kishan et al. (2015)	Ant colony optimization (ACO)	Chooses optimal path which has no malevolent vehicles	Choose the optimal route for safe performance of the message	Updating of trust and pheromone is overhead
Karan Verma et al. (2013)	IP check	Become aware of DoS by means of analyzing the IP address	Preserve on the way to IP spoofing of content with and preserves bandwidth	It only detects DoS
Hsiao et al. (2011)	FastAuth and SelAuth	Prevents signature flooding	Fast authentication with short signatures	In case if a packet is misplaced, complicated to authenticate

are responsible for several categories of assaults in which attacks on the availability of societal guidance are of full-scale difficulty. One of the essential security issues that occurred in VANET is denial of service attack. Here, a brief survey is provided on denial of service attack and also discussed the various solutions to detect and prevent DoS attacks in vehicular ad hoc environment. Finally, the comparison is made between the various solutions for the DoS attack in VANET.

References

1. Qian Y, Moayeri N (2008) Design of secure and application-oriented VANETs. In: VTC spring 2008-IEEE vehicular technology conference. IEEE, pp 2794–2799
2. Jakubiak J, Koucheryavy Y (2008) State of the art and research challenges for VANETs. In: 2008 5th IEEE consumer communications and networking conference. IEEE, pp 912–916
3. Liu M, Zeng X, Tao M, Cheng Z (2018) A new clustering protocol for vehicular networks in ITS. In: International conference on communications and networking in China. Springer, Cham, pp 615–624
4. Bariah L, Shehada D, Salahat E, Yeun CY (2015) Recent advances in VANET security: a survey. In: 2015 IEEE 82nd vehicular technology conference (VTC2015-Fall). IEEE, pp 1–7
5. Singh A, Sharma P (2015) A novel mechanism for detecting DOS attack in VANET using enhanced attacked packet detection algorithm (EAPDA). In: 2015 2nd international conference on recent advances in engineering and computational sciences (RAECS). IEEE, pp 1–5
6. Saranya R, Kannan SS, Prathap N (2015) A survey for restricting the DDOS traffic flooding and worm attacks in Internet. In: 2015 International conference on applied and theoretical computing and communication technology (iCATccT). IEEE, pp 251–256
7. Mejri MN, Achir N, Hamdi M (2016) A new security games based reaction algorithm against DOS attacks in VANETs. In: 2016 13th IEEE annual consumer communications and networking conference (CCNC). IEEE, pp 837–840
8. Kumaresan G, Macriga TA (2017) Group key authentication scheme for vanet intrusion detection (GKAVIN). Wireless Netw 23(3):935–945

9. Jamil F, Javaid A, Umer T, Rehmani MH (2017) A comprehensive survey of network coding in vehicular ad-hoc networks. Wireless Netw 23(8):2395–2414
10. Faghihniya MJ, Hosseini SM, Tahmasebi M (2017) Security upgrade against RREQ flooding attack by using balance index on vehicular ad hoc network. Wireless Netw 23(6):1863–1874
11. Ibrahim S, Hamdy M, Shaaban E (2015) A proposed security service set for VANET SOA. In: 2015 IEEE seventh international conference on intelligent computing and information systems (ICICIS). IEEE, pp 649–653
12. Dong P, Du X, Zhang H, Xu T (2016) A detection method for a novel DDoS attack against SDN controllers by vast new low-traffic flows. In: 2016 IEEE international conference on communications (ICC). IEEE, pp 1–6
13. Rehman SU, Manickam S (2015) Rule-based mechanism to detect denial of service (DoS) attacks on duplicate address detection process in IPv6 link local communication. In: 2015 4th international conference on reliability, Infocom technologies and optimization (ICRITO) (Trends and Future Directions). IEEE, pp 1–6
14. Erskine SK, Elleithy KM (2019) Real-time detection of DoS attacks in IEEE 802.11 p Using Fog computing for a secure intelligent vehicular network. Electronics 8(7):776
15. Guleria C, Verma HK (2018) Improved detection and mitigation of DDoS attack in vehicular ad hoc network. In: 2018 4th international conference on computing communication and automation (ICCCA). IEEE, pp 1–4

Cloud-Based ETL and AD Impression Prediction and Insight Visualization Using Tableau

G. Tamilmani and K. Rajathi

Abstract In any small-scale or large-scale business, digital marketing plays a crucial role in the brand reach and success. But, the cost of digital marketing has increased significantly over the years as Internet grew from being a luxury a few years ago to a common commodity today. When an ad campaign is started, even though ad targeting technologies are already available, there is a significant amount of wastage of advertisement budget, as many users just look over the ad and never click. The idea behind this project is to help the businesses to invest wisely in digital marketing and target their audiences at a more personal level. This project will help businesses to identify the proper media platform for running the campaign for specific people. We use Google analytics to gather the data of the previous ad campaigns and develop predictive model which predicts the probability of the user clicking on an advertisement in a particular Web site. As the database contains data of several users, it is large and difficult to perform operations using the local resources; we are relying on cloud-based tools like snowflake for data warehousing and Talend to perform ETL tasks.

Keywords Cloud warehousing · ETL · Machine learning · Digital marketing

1 Introduction

This project aims to help businesses to spend their marketing budget in a more efficient way. This project mainly relies on cloud warehousing, advanced analytics, and machine learning.

With increasing digital marketing costs, it is high time that we not just target specific people but also target the platform where we post advertisements. However, conventional marketing solutions assist us in identifying the correct people, but they do not recommend the appropriate medium in which the consumer is most likely to

G. Tamilmani (✉) · K. Rajathi
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai, Tamil Nadu, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_28

click on the advertisement and make a purchase. We started this project with a goal to solve this problem.

2 Literature Survey

Digital channels of communication for promoting information about products in the catalog, studies have used the term "digital marketing" as a collective term representing all of the content-based digital platforms and device. It has already been established that the majority of customers use digital channels while searching and evaluating the products and services, yet the capability of digital channels in stimulating the need of products buying largely remains unknown.

It has already been established that the majority of customers use digital channels while searching and evaluating the products and services, yet the capability of digital channels in stimulating the need to increases sales largely remains unknown [1]. Use of hypothesized digital marketing communication affected the important activities and key decision variables representing the entire buying decision process. A positive effect was said to be made when a statistically significant higher-than-average value was found associated with every stage.

Widely varying retailing formats have emerged to satisfy the demands of consumers across diverse physical, geographical, purchase, and need situations. In most instances, these retail formats take the three form of physical stationary retailers, with the result that the division of labor between providers of consumer goods (manufacturers), and those engaged in distribution and sales (wholesalers and retailers) has been rather stable over a long history [2–4]. Online retail platforms have been successfully established pivot end to claim the primary customer interface in the retailing value chain. Examples are Amazon marketplace, Alibaba, eBay, and JD. These platforms employ digital technologies as intermediaries in the exchange of products and services between buyers and sellers digital technologies simplify customers' access to and retrieval of comprehensive product information from various sources (e.g., online product descriptions, reviews, prices). Further, digital technologies can facilitate the analysis of relevant data (e.g., through machine learning algorithms) and the transformation of these data into insightful information.

Authors [5, 6] Establish a connection between the layout of objects in a display and the amount of attention devoted to each of these objects. The research has implications for catalog page layout and visual merchandising. The applications, calculations, depictions, and references to other related disseminations are all distributed using a continuous univariate distribution [7, 8]. It is organized for easy navigation and quick reference.

The project is basically used for finding best suitable media platform to promote the product or new lucrative deal on new brand. It projects deal with the prediction model to predict best media channel or online streaming about which can bring the maximum cash flow and generate the maximum to increase the sales. In this, review

described about the project strategy and idea behind project considering to generate maximum cash flow for an e-commerce site.

3 Proposed System

The proposed system uses the previous ad campaign data to suggest the digital platform for different users. This helps in effective ad campaigns, both in terms of cost and conversion rate. This solution is economically feasible, as it is just a one-time investment while building the predictive model. Later, the price incurred is only for the ad campaign but not for using the solution itself.

This solution requires minimum training and could be easily operated by anyone, as this is quite user-friendly. Hence, this solution is feasible from operations stand point.

Figure 1 shows architecture which depicts the workflow of project. It shows the data origin from Google analytics and its storage in snowflake warehouse. Snows are used to interact with the warehouse. The ETL and visualization tools are used to find the insights from within data which could be helpful in building the predictive model.

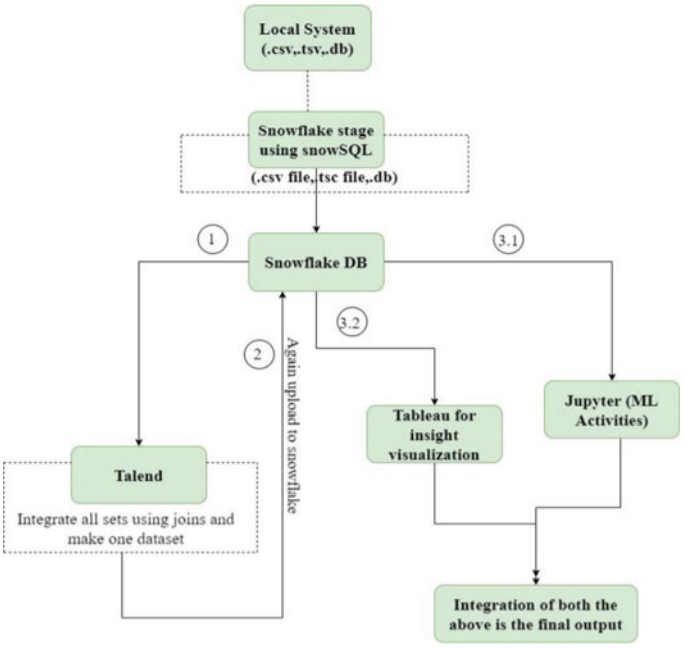


Fig. 1 Architecture diagram

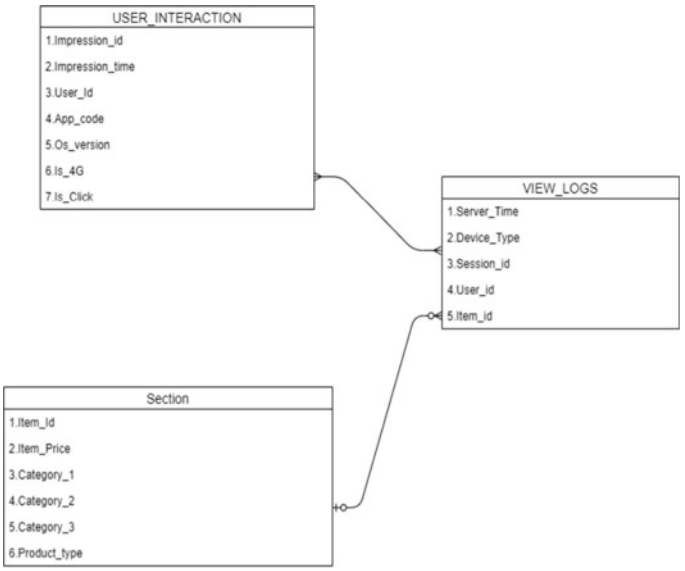


Fig. 2 ER diagram for interconnected data

Figure 2 shows, i.e., entity-relationship diagram clearly shows how our data are interconnected with foreign and primary keys. The data are relational, and the necessary data are extracted by performing proper joins on the data as required.

The sequence diagram Figure 3 shows object interactions arranged in time sequence which is used in the project. It shows how the different objects of the project are related while interaction from preprocessing data to predict from the data. It shows how each function in the project takes place.

Figure 4, the collaboration diagram is an illustration of the relationships and interactions among software objects in the unified modeling language (UML). This diagram can be used to portray the dynamic behavior of a particular use case and define the role of each object.

In this, the output of one phase is the input for another phase. The output of the data extraction phase is the input of the ETL phase. The output of ETL phase is the input for insight generation phase. The output of insight generation phase is the input of ML model building phase.

4 Result and Discussion

Efficiency of the proposed system: The efficiency of the proposed system is 8–10% than the existing system. The prediction generated through the proposed system is more precise and accurate about the data. Analysis of data using the visualization technique in proposed system makes it more easier to analyze the data.

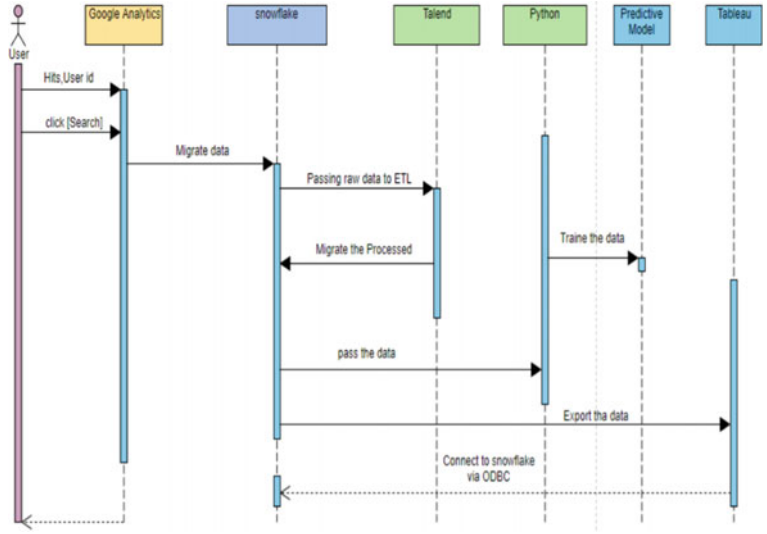


Fig. 3 Sequence diagram

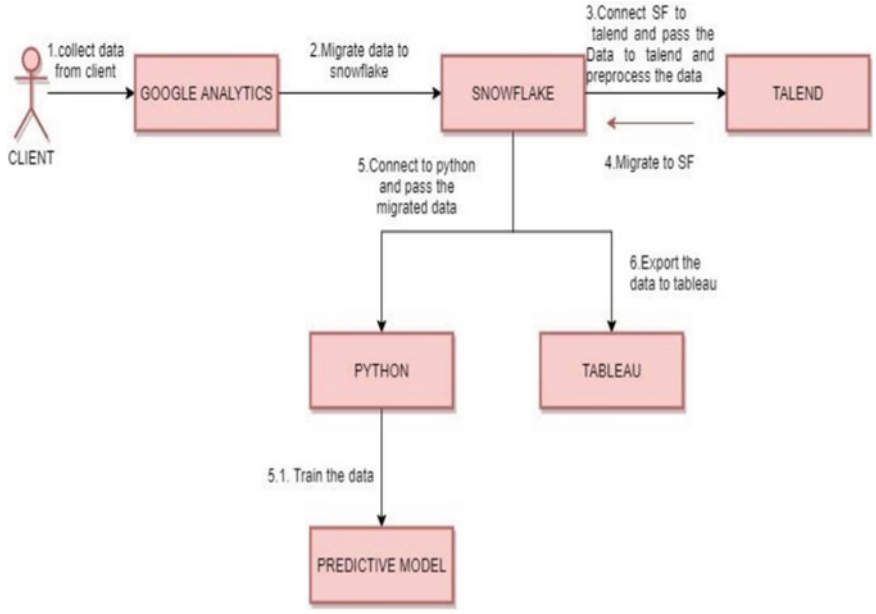


Fig. 4 Collaboration diagram

Comparison of existing and proposed system: The proposed system uses a prediction model for increasing the conversion rate based on previous ad campaign data instead of selecting random channel based on ambiguous data. There is no guarantee for the user to be effected by the advertisement posed by the company. Precision of the proposed system prediction model is far better than existing system since it is able to make prediction with 89.9% accuracy.

5 Conclusion and Future Enhancement

This project concentrates on minimizing the costs that are required for digital marketing. This solution helps to target the ad campaigns in a more personal level to each user, in turn increasing the conversion rate for the ad campaign. By implementing this entire solution in cloud, there is now no need to worry about the local computational power.

Future enhancements: This project is designed and developed for using a particular prediction model for generating the results. In future enhancements, the project will be generating its own model based on data to increase its accuracy. The project will be using neural model for the better prediction in future.

References

1. Dahiya R, Gayatri (2018) A research paper on digital marketing communication consumer buying decision process: an empirical study in the Indian passenger car market. J Glob Mark
2. Alba JW, Hutchinson JW (1987) Dimensions of consumer expertise. J Consumer Res
3. Babin WRD, Barry J, Griffin M (1994) Work and/or fun: measuring hedonic and utilitarian shopping value. J Consumer Res
4. Moe WW, Fader PS (2000) Capturing evolving visit behavior in clickstream data. Wharton Marketing Department Working Paper
5. Janiszewski C (2018) The influence of display characteristics on visual exploratory search behavior. J Consumer Res
6. Jarboe GR, McDaniel CD (1987) A profile of browsers in regional shopping malls. J Acad Mark Sci
7. Lenk PJ, Rao AG (1995) Transition times: distributions arising 27 from time heterogeneous Poisson processes. J Manag Sci
8. Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, vol 1, 2nd ed. Wiley, New York
9. Reinartz W, Wiegand N, Imschloss M (2018) The impact of digital transformation on retailing value chain
10. Bloch DLS, Peter H, Ridgway NM (1986) Consumer search: an extended framework. J Consumer Res
11. Punj GN, Staelin R (1983) A model of consumer information search behavior for new automobiles. J Consumer Res
12. Hirschman EC (1984) Experience seeking: a subjectivist perspective of consumption. J Bus Res

Object Recognition and Multiple Image Stitching

R. Kavitha and P. Subha

Abstract Images are used in diverse engineering works and scientific research. But, every image has a particular resolution and size. Image stitching is one such possible way which can increase the resolution and size of image. This is a technique through which we can join multiple images which has some common key point and can make the image a wide-angle image. It could be one of the inexpensive ways to determine the wide-angle view. Cases like landslides, earthquakes, or any natural disaster could affect a wide area, but when we go for survey or estimation of damage, it can be an expensive method to arrange helicopters and gather the view of the place. Their image stitching ought to play an inexpensive and essential role. This report also focuses on the improved and efficient methods of traditional image stitching and reducing the errors. The old-style image stitching results are based on methods like feature point extraction and it results to some extent of misrepresentation errors, which is frequently occurred due to extended image arrangements. To provide high-quality image, an improved algorithm is proposed. The test aftereffects of improved algorithm have demonstrated that we can quicken the productivity of picture merging process, and furthermore lessen the all-encompassing distortion blunders. Lastly, object recognition can make the stitched information more informative by labeling. Using image stitching and object recognition, we can get multiple achievements helping us to get high-resolution panoramic view, increasing our possibility to get a wider view of any. These UAVs will help in getting the view from top in cases like earthquakes and landslides loss estimation. Using Internet of Things and machine learning, it can be achieved very efficiently.

Keywords Stitching · Object recognition · SIFT feature extraction · Panorama · IoT · Machine learning

R. Kavitha (✉)
Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India
e-mail: rkavitha@veltech.edu.in

P. Subha
Department of IT, Sri Sai Ram Institute of Technology, Chennai, Tamil Nadu, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_29

1 Introduction

1.1 Multiple Image Stitching

Getting an unobstructed and wide-angle view in all directions is called panoramic view. Many of the times, when we need to get a wide-angle view of any place, we choose for panoramic mode in our smartphones. But, in Internet of Things and machine learning, image stitching is a method used to get the wide-angle view when we deal with robotic applications and machines, or unmanned aerial vehicles. It is used to get the wide-angle high-resolution image programmed as per requirements.

Multiple stitching of images is a procedure to connect a succession of pictures with one another. The images have overlapping regions commonly converting into a consistent picture. The camera has good resolution and clarity.

Image registration step is a substantial process in the multiple stitching of images. The nature of multiple stitching of images incredibly relies upon the accuracy of image deployment. With respect to the image registration, the multiple stitching of images and their calculations are partitioned into two classes, namely feature based and region related. Region-related image registration class anticipates the relationship of the same dimension blocks between the input and the reference image and records their equivalence value. If in any case the picture is turned or resized, this strategy cannot not bring an ideal outcome. In any such case, the image textures were excessively strong or weak; the outcome would show massive mistakes. Feature-based image registration strategy uses mathematical models to locate the abstract description features of the valuable pixel data by contrasting the description features with discover the correspondence association between the input image and the reference image.

Numerous stitching of pictures is a huge development in the image preparing area. The purpose of numerous picture sewing is to change different source pictures with territories covering each other to tie together in a comparative arrange system through change framework. Thusly, it is basic to pick a reference to organize structure. The customary consolidating measure constructs all-encompassing pictures from requested picture groupings, sewing little by little from left to right. The principal image of the plan is picked as the reference picture. As such, select the air conditioner companying sewing result as the accompanying new reference picture in the customary consolidating measure. The customary consolidating measure cumulates the coordinating bungles of each sewing technique in the reference picture, which would be really misshaped; when the course of action of picture groupings is huge, it impacts the idea of all-encompassing result. The improved method proposed in this report initially executes picture enlistment for each and every neighboring picture in the progression and registers the change framework between the adjoining pictures and subsequently took the middle picture of the gathering as the reference picture. As shown by the change lattices of each and every close by picture, the improved procedure can comprehend picture wherever in the progression changes to the arrange space of the reference picture. Likewise, all photos in the plan can be

bound together to a comparative encourage structure in the wake of encountering all sewing structures. The exploratory results show that the improved technique lessens the mutilation mistakes of scene and extras the hour of sewing measure; furthermore, it redesigns the idea of the sewing result. Object recognition is an advanced mechanism to understand the objects present near to see. Years back if we observe, such method could be identified as advanced virtual artificial eye, but with development of technology, we figure out that the algorithm, machinery computation power, and advanced datasets have made easier to devise optimized method for object detection.

Joining image stitching and object recognition in Internet of Things, with a proficient and quick handling hearty processor, we can accomplish different applications in the field of following, observation and security, and so on.

2 Literature Review

Multiple image edging is the combining of many images with overlapping fields to produce a segmented panoramic image or high-definition image. It is used in different type of industries based on the purpose [1, 2].

UAV imagery interpretation [3], UAV are used in soil monitoring to classify the different types of land and calculate vegetation indices using multispectral imaging [4].

UAV and image stitching utilizing machine learning are assuming imperative job to assemble a brilliant strong waste location framework which utilizes the movement of an unmanned aerial [5, 6].

Moreover, UAV fast stitching technique of remote sensing data, for disaster prevention and mitigation departments, has become a more and more important first-choice information source of first-hand disaster information extraction [7, 8].

Object detection is mainly based on identifying real-world objects. In this paper, various algorithms are used for object detection such as face mask detection, iris detection, fingerprint detection, texture detection using MATLAB to detect objects for video surveillance applications. Skin detection algorithm fails when multiple objects are present. Other challenges are change illumination, detecting shadows, modeling of non-static background, and difficult to classify a fixed foreground and background [9].

Object detection is mainly based on identifying real-world objects. Various algorithms for video surveillance applications are used are discussed in this paper. Parameters such as detection accuracy, RGB Euclidian threshold ‘T’ in target detection, Y, Cb, and Cr in skin detection have been simulated and implemented to improve the efficiency of the algorithms. Individuals in upstanding position can be identified utilizing the cardboard model. In the event that individuals are in various stances, or are slithering and climbing, it gets testing. Skin discovery calculation bombs when different items are available. Different difficulties are change enlightenment, identifying shadows, displaying of non-static foundation, and hard to group a fixed front ground and foundation [10].

3 Existing Work

The traditional image stitching algorithm has proved to be slow and less efficient with distortion errors. The traditional algorithm is based on the SIFT feature points extraction. While compositing an all panoramic outcome utilizing a long image sequence, the output panoramic image gets distorted.

3.1 Image Registration

As proposed by Brown and Lowe in 2004, they concocted scale-invariant highlights extraction computation (SIFT) and reliably proposed the numerous images sewing measure and proceeded till 2007. The SIFT computation identifies the image includes rapidly and gives invariant property when the image is rotated, resized, or enlightened [11].

Enormous scale factors compare to the diagrams of the image features and little scale factors relate to the details of the picture features (Fig. 1).

Each examining point is contrasted and the eight neighborhood points of a similar scale factor and 18 area purposes of the neighboring scale factors. The recognized outrageous focuses are chosen as the up-and-comer key focuses.

Next need to guarantee the revolution trademark, turn the arrange hub to the direction of the central issue and afterward partition the locale around the central issue into 474 subregions. Every subregion has 8-measurement vector of direction data. Accordingly, there is $4 * 4 * 8 = 128$, measurement vector in the central issue descriptor. Finally, we standardize the central issue descriptor to wipe out the impact of the enlightenment change.

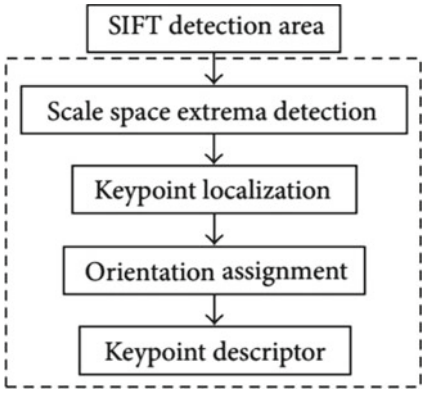


Fig. 1 SIFT feature extraction step

4 Proposed Methodology

4.1 Ground Station for Distributed Processing

Ground station is a fully working and managed service that lets us to control distributed processing of multiple cameras such as giving run command for executing programs, processing of stored image data. For example, unmanned aerial vehicles are used for a wide range of applications, like monitoring and surveillance, image processing, surface imaging, communications, survey and rescue in natural disaster, and video broadcasts. Ground stations form the core of swarm of UAV networks. Jetson nano is being used as ground station in this application where all data coming from drones are going to be stored and get stitched each other for video generation. YOLO V3 object detection model is going to be performed for object detection in video on ground station.

4.2 Image Acquisition

Image acquisition is the taking of images from user as input using any hardware source such as cameras. It is the first step of any image processing application. In proposed work, consider input using GoPro camera. GoPro cameras are the high-definition cameras which uses image-stabilization factor to prevent the image from producing blurred image (Fig. 2).

For unmanned aerial vehicles, it can prove itself as a successful application, when in flight mode vibration produced upon using simple cameras could not produce clear

Fig. 2 Image acquisition using GoPro camera

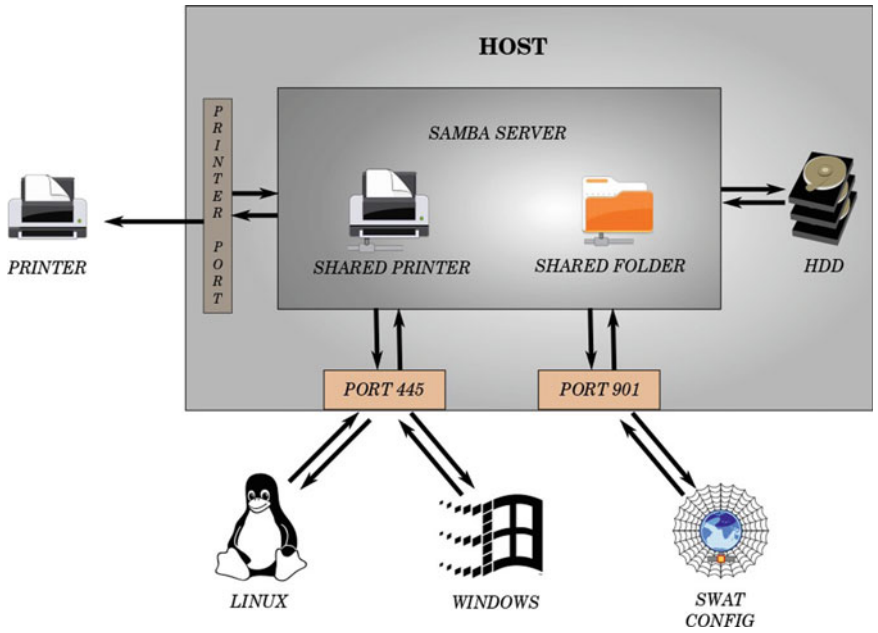


Fig. 3 Linux samba share mechanism

image. GoPro cameras are also enabled with point-of-view photography shoot. For this application, we need a bright light condition for our image stitching algorithm to detect the descriptors which are possible with a good quality camera with good resolution, and this could be achieved with GoPro cameras. Installing GoPro camera could result into a best result. However, user's choice depends upon their use (Fig. 3).

4.3 Linux Samba Share

A Linux Samba file server allows file sharing across different operating systems over a network. It allows you to access your desktop files from a system and share files across cross-platform such as Windows and macOS users. It can be used for sharing the images file taken by multiple cameras to store in ground station for further processing.

The Samba suite incorporates customer instruments that permit clients on a Unix framework to get to envelopes and printers that Windows frameworks and Samba workers offer on the organization. The Samba suite depends on Unix daemons that give shared assets called shares or services 7 to SMB customers on the organization.

4.4 YOLO Dataset

YOLO or You Look Only Once is a state-of-art dataset which is used for real-time applications such as object detection system. Configuring the object recognition in this application, this dataset was a suitable choice. It processes images at 30 FPS. It is tremendously wild and precise. Moreover, no retraining is required for this dataset and have tradeoff between accuracy and speed.

4.5 ImageAI

ImageAI is a Python framework which classes and functions to perform image object detection and extraction in an easier way.

4.6 Improved Stitching Method

The proposed calculation doesn't distinguish the component focuses for the whole picture territory; however, it likewise uses the factual enrollment information of each and every close by picture to deal with the grid. Subsequently, the improved system decreases the SIFT include recognition region of the reference picture, and besides constructs the amount of coordinating element focuses and improves the capability of sewing time. Similarly, it fastens consecutively from left to right independently in the customary procedure; the improved methodology picks the middle picture as the reference picture and picks the accompanying information picture dependent on number of planning centers and completes the image arrangement sewing measure from focus position continuously developing to the different sides. In the proposed sewing technique, the most outrageous assembled coordinating bumbles are resolved from the middle picture to the various sides of the image grouping, lessened to the half level of the customary methodology.

4.7 Algorithm

The improved image stitching process.

Input n well-ordered picture sequence S (S_0, S_1, S_2, S_3, ….$S_{n-1)}$, those have overlapping regions.

1. Perform image registrations for nearby images.

   (a) Consider the images are I_i and I_{i-1}, $i \in [1, n-1]$, use k-nearest neighbor method to find matching feature points from Ii and $Ii - 1$ considering the

Fig. 4 Object recognition showing labeled boxes

base Manhattan distance and afterward protect the coordinating component focuses information. Save the pictures in an exhibit using fList and nList.

(b) Process the network that can be moved from the pursuit space of the picture I_{i-1} to the inquiry space of the picture Ii and afterward protect the relative lattice in the exhibit HList.

(c) Repeat step (i).

2. Take Ik as the reference image. K is the middle index of the set I.
3. Process the matrix H between the new input image and the reference image.
4. The algorithm is computed to optimize the matrix array H.
5. Transform the new input image using H matrix (Fig. 4).
6. Compute the affine transformation result and process the step (5) and the reference image.
7. Add the input image and the reference image.
8. Panoramic result is produced by repeating step 2.

Output
Panoramic image

4.8 Object Recognition Using *ImageAI*

Image recognition is the part of PC vision, where outwardly perceptible picture outlines in video can be distinguished, restricted, and perceived by PCs. A solitary static picture in a video is known as a video outline, and in the vast majority of recordings, the scope of edge every second or fps lies between 20 and 32. With the help of Python image library which includes ImageAI, which comes in-built with

self-contained deep learning and computer vision capabilities. It provides classes and functions to perform image detection and extraction.

4.9 Comparison with Other Models

In our algorithm, we are using YOLOv3 which is supported by ImageAI. This object recognition is applied to the later part of the image sequence conversion in which the all-stitched images are converted into a single video output, and it is fed as input, and we get the output as the labeled video in which has all the objects are detected, and they are labeled with boxes with their respective identity named on top of box with their verification percentage (percentage stating the chance of the object being real as labeled). For this object recognition, as they are a video input, primarily, they are needed to be converted into frames. Multiple frames are processed and recognized one after another by the algorithm.

5 Conclusion and Future Enhancement

The algorithm based on machine learning is used to stitch the images, share files along the server, and perform object recognition using NVIDIA Jetson Nano and GoPro cameras through Python programming language. It can be used for various applications such as surveillance, tracking and damage estimation, producing high-resolution images, high-definition images, estimation of damage using unmanned aerial vehicles in earthquakes, landslides, medical imaging, video conferencing, and many more. It can be used for basis of producing 3D image stitching and applications.

If equipped with more powerful processors, it can produce more efficient results. Image stitching can be improved more by adding image stabilization. And, the program can be made more fast if executed in cloud using cloud technology and Internet of Things.

References

1. Casbee DW, Li S-M et al (2005) Forest fire monitoring with multi- ple small UAVs. In: Proceedings of American control conference pp 3530–3535
2. Niethammer U, Rothrnund S, Joswig M (2009) UAV-based remote sensing of the slow moving landslide Super-Sauze. Landslide Processes, CERG Editions, Strasbourg, vol 352, pp 69–74
3. Lin 352, Tao H, Wang Y, Huang Z (2010) Practical application of un-manned aerial vehicles for mountain hazards survey. In: 8th international conference on Geoinformatics, Beijing, pp 1–5. https://doi.org/10.1109/GEOINFORMAT-ICS.2010.5567777

4. De Biasio M, Arnold T, Leitner R, McGunnigle G, Meester R (2010) UAV-based environmental monitoring using multi-spectral imaging. In: Proceedings of SPIE. The international society for optical engineering https://doi.org/10.1117/12.864470
5. Bang SK, Kim H, Kim H (2017) UAV-based auto-matic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching. Autom Constr 84:70–80. https://doi.org/10.1016/j.autcon.2017.08.031
6. Ming, L, Huan L, Xinyan Z (2012) Approach to fast mosaic UAV images for disaster emergency. J Catastrophology 27(3):139–144
7. Ham Y, Han KK, Lin JJ et al (2016) Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs), a review of related works. Vis Eng 4(1) https://doi.org/10.1186/s40327-015-0029-z
8. Raghunandan A, Raghav P, Aradhya HR (2018) Object detection algorithms for video surveillance applications. In: International conference on communication and signal processing
9. Guodong C, Xia Z, Sun R, Wang Z, Sun (2018) A learning algorithm for model-based object detection: Sensor Review ISSN: 0260-2288, 33(1):25–39. https://doi.org/10.1108/02602281311294324
10. Qu Z, Lin S, Ju F-R, Liu L, Sun L (2015) The improved algorithm of fast panorama stitching for image sequence and reducing the distortion errors, Hindawi Publishing Corporation Mathematical Problems in Engineering, vol 2015, pp 12
11. Sujeetha R, Mishra V (2019) Object Detection and tracking using tensor flow. Int J Recent Technol Eng (IJRTE) 8

An Efficient Approach for Analysis and Segmentation of Plants Diseases Using Machine Learning

R. Umamaheswari and S. Uma

Abstract The farming assumes a predominant function in the development of the nation's economy. Atmosphere and other natural changes have become a significant danger in the horticulture field. Farmers are really facing lot of loss in crop production due to many reasons; one of the major problems for the above issue is crop diseases. This is due to enlightenment about the disease and pesticides or insecticides available in order to control the disease. AI (ML) is a fundamental methodology for accomplishing reasonable and powerful answers for this issue. Recognizable proof of sickness inside the plant may be extremely basic for better cultivating outcomes. In this paper, backpropagation neural network and neural network classifier (PatternNet) are used for achieving the result.

Keywords Feature extraction · Segmentation process · Backpropagation neural network · Network classifier

1 Introduction

Agriculture plays a vital role in our country. In recent years, many researches are happening in the agriculture sector like predicting plant disease, predicting the leaf diseases, predicting the yield. One of the key resources in efficient management of the crop is to predict the diseases at the earlier stage and to provide accurate results and on-time solutions to the farmers. Generally, in any sector, prediction is a very difficult task as it includes various features. Identifying the disease present in the plant leaf for better agricultural productivity is one of the difficult tasks. Plant diseases are primary cause for enormous damage and economic losses in the agricultural field. Identification of ailment inside the plant might be very useful for producing the better farming results. It is conceivable to build up a self-governing framework for illness characterization of yields.

R. Umamaheswari (✉) · S. Uma
Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi, Chennai, Tamilnadu 600062, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_30

Identification of the diseases will help the farmers to identify the faults in the plant leaves which will lead to precautionary methods for preserving the other plants. In this paper, with the help of the image processing and pattern analysis methods, based on the symptoms of the particular diseases, the damaged parts of the leaves are captured through the image processing which will undergo further processing to predict name of the diseases of that particular plants. Subsequently, after all the process has been completed, the name of the diseases that affected the plant can be predicted by considering all the relevant factors, and the recommendation will be given to the farmers in order to take the corrective prevention actions.

2 Related Works

It discusses about analysis of various literature surveys. In the recent years, various machine learning algorithms and techniques have been in the hand of the researcher to study about the plant leaf diseases; some of the reviews on them are given below. Author [1] presents about reorganization of powdery mildew spores diseases which occur more often in the plants. But, it can be identified very easily since symptoms of these diseases are more peculiar, and the diseases will often occur on upper surface of the plants. Here, author used an intelligent spore image sequence capture device approach to capture the damaged leaf parts. To this system, when the damaged leaf part has been passed as input, it would be naturally started up and checks blue print, and measures has been satisfied by the input. Finally, the output will be passed to the pre-specified network.

This author [2] presents about exposure of diseases that can occur in the plants with the help of remote sensing images. For creating the training datasets, images of healthy as well as damaged or diseased leaves are collected. From the trained samples, the maximum threshold values are extracted which are used to predict the incoming input image. The images that are collected using the remote sensor are often compared with the gained threshold value, by extracting the RGB values that lie in the image.

Author [3] presented about detection of yellow vein mosaic disease in okra plants. Yellow vein mosaic disease is viral diseases which are spread by insects. Here, the author used the leaf vein extraction and naive Bayesian classifier in order to predict the diseases. Author [4] used the vigilant edge identification methods in order to recognize the infected part of the plant, and the information regarding the diseases will be sent as an alert message to the farmers via the message transferring technology that lies in the MATLAB.

3 Some Effects of Bacterial/Fungal/Viral Diseases in Plants

Suppose if the portion of the plant is found to be sick, the approach consists of splitting the risk region (ROI) and deciding the illness. Throughout this question, the groups of leaf pain, including blight spot, necrosis, and late spot are listed. Many of these infections are triggered by fungal infections or bacterial infections. The bacterial leaf location is also established. Such as strawberry, peach are mostly located in a bacterial leaf spot that prevents from being stoned. Dark streaks or black stains appear from this suffering on the different portion of the stems. A sign of this disorder is often yellow halos. The scale of the place in person is anomalous. Spots may occur on the specific item, and when such points are concentrated on some part of the vine, the results would be obtained on destroying the process with the aid of this disease. In fact, muddy and foul formations lead to bacterial ailment development. Bacterial leaf spot will spread very quickly in these formations. The elderly leaves are mainly bacterial in nature, but they may also kill the tissue of the young leaves.

Viral illness is attributable to viruses, because, since the infection is intracellular, it occurs indoors. Viral infections become really challenging to grasp hopefully. In all the plants such as leafs, roots, stem, and others, viruses may be successful. Abnormal patterns are seen in the region infected by the virus in the green and yellow patterns. The lifestyle or additives of the plant that are affected by the viral disease can be much smaller. It affects productivity and different elements at the same time. For a particular case, consuming a part of the blood is often the number one sign of this disease. Different viruses may be severe in assessment to the alternative styles of the disorder, because of the reality every virus if now not nicely cured give upward thrust to sort of the virus, so it is crucial for well-timed of those sickness. Owing to fungi or fungal species, fungal diseases seize the area. This grows with wind and water, one of the properties of fungi. Gray inexperienced points are on the flora's leaf, and they begin to stretch out closer to the outside position of the leaf if they are no longer properly cured. The indications, scabs, and manifestations of fungal infection are the number one effects. The assault of fungal diseases on the leaves of the plants prevents the yellowness of the plants.

4 Proposed System

To identify the disease present in the plant leaf for better agricultural productivity, helping the farmers to identify the faults in the plant leaves will lead to precautionary methods for preserving the other plants. After identifying the key points, the following step is to come back up with a descriptor for the function targeted at every key point. This descriptor is the illustration used to take a look at the talents in specific photographs to look inside the event that they are wholesome or no longer. Distinguishing proof of sickness inside the plant may be fundamental for better cultivating outcomes. The key rationale is to accurately distinguish leaf ailments with the

utilization of neural organizations. Since the component extraction approach seems to a basic job, here, feature extraction techniques such as the speeded up robust features (SURF) is used for extracting the factor.

Speeded up robust features (SURF) descriptor scheme advanced by way of beetle became designed as a green alternative to SIFT. It is a lot quicker and extra strong as compared to SIFT. Its fundamental idea is to approximate the second one-order Gaussian derivatives in a green manner with the assist of essential pictures for the utilization as set off field filters. This concept uses sum of 2D Haar wavelet responses in order to round of the focal value with the useful resource of vital picture.

- Using deep neural networks for classification purpose.
- To determine the disease location in the plant leaf using image

Then, the input image in RGB form is converted to hue saturation value (HSV) form for the segmentation process. The leaves which have the ailment are then classified based on the premise of their similarity with the training dataset of disease samples formerly defined by using the feature extraction techniques. Here, the implementation is done with backpropagation neural network, and neural network classifier (PatternNet) is used for which results in displaying the name of the disease present on the plant leaf (Fig. 1).

This figure shows the architecture data flowing through different stages in order to give the output. Digital gadgets including digital camera or smart phones are used to take pictures of plant leaf and are used for splitting up and grading the nearby sickness. To discover the affected area efficiently, it is required that a pictures containing impurities in order to compare other image processing techniques in the photograph entry.

5 Methodologies

5.1 Image Acquisition

The processing of photographs is to produce those forms of sample pictures. It is really important to use the dataset to cover the majority of the various types of inputs and have an excellent tool for every problem. The repository contains various images. In the collection of matching rules (with the disease leaf samples in the dataset), a leaf may be added and its shape is understandable, and the classifier provides for the disease prediction for this random picture (Fig. 2).

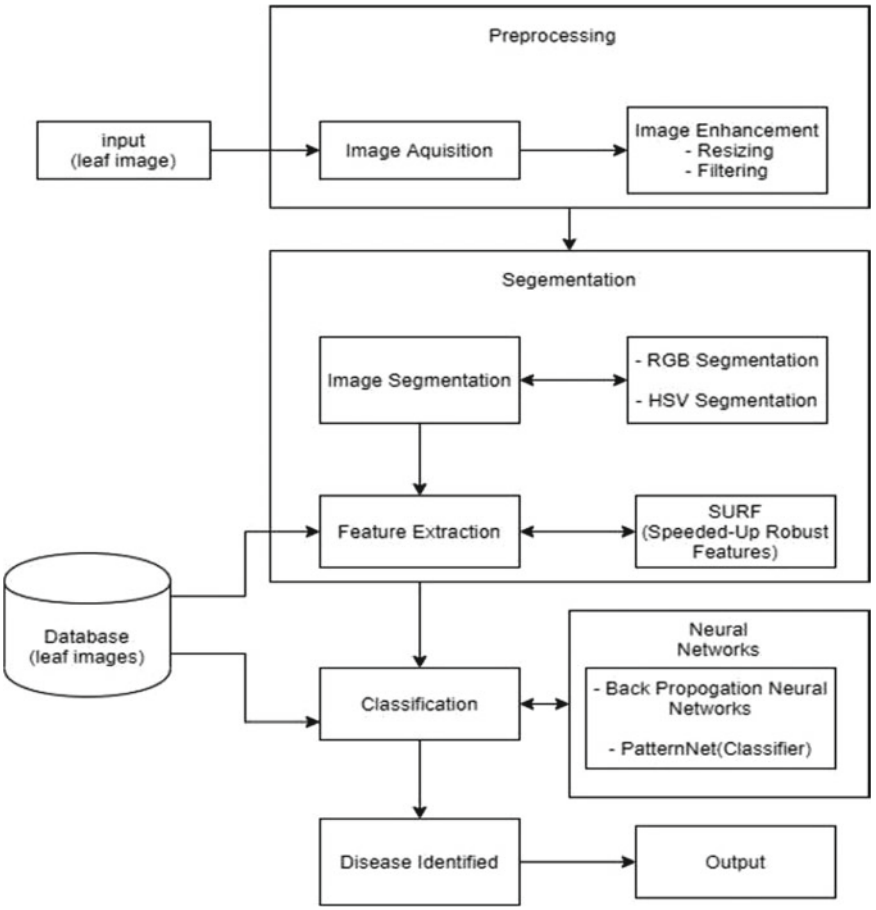


Fig. 1 Architecture diagram

5.2 Image Preprocessing and Enhancement

Photography, it is the second segment. By rising input noise from MATLAB, the pixel values are graded more; there is a discount on spots and an increase in assessment. This image planning and refinement are attributed to the fact that it is easy to segment the hobby position, while the photograph is preprocessed. The classifier used will implicitly functions well without any impurities for preprocessed images. Within the preprocessed image, the pixel values are often changed. For enhancement of the image, MATLAB offers a variety of various filters.

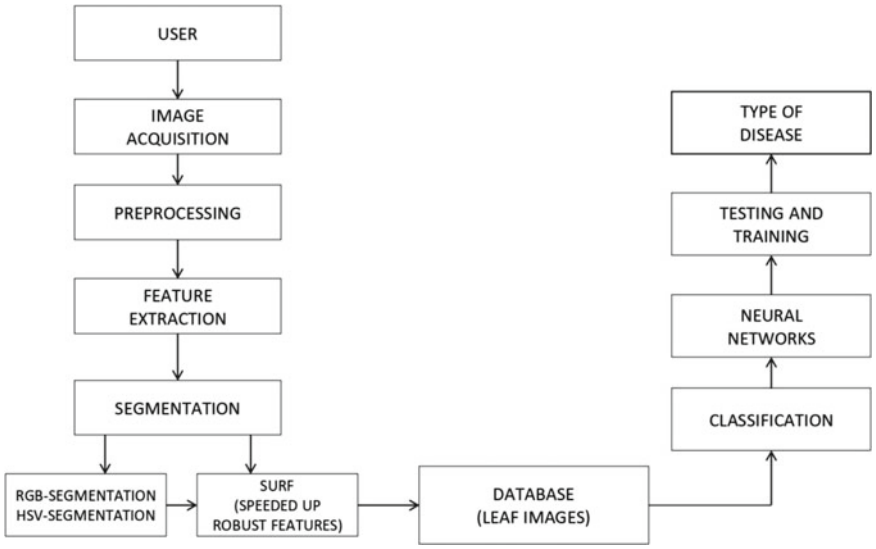


Fig. 2 Flow of methodologies used in proposed work

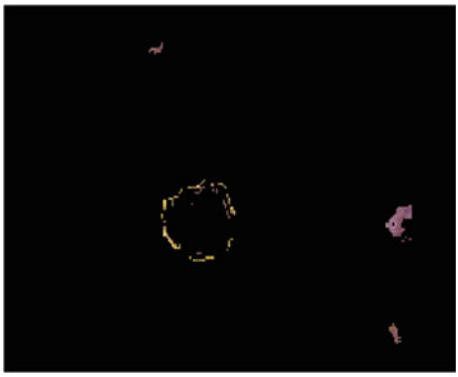
5.3 Image Segmentation

To make a detailed analysis of the leaves, the subsequent steps are performed and the codes are written in MATLAB 7.50 version. First, read the image, convert into RGB to HSV, set image as h = hsv(1); s = hsv(2); v = hsv(3), then in order to calculate number of pixels and plots that are present in a histogram, determine the location of black and white pixels, to grab estimate of all pixels for each color bin. The idea of primary colors is primarily based totally on the truth that any color can be created through combining three colors, white, and black. However, with the resource of reflected photograph (via paint), there are more that definition where in set of colors are the "number one shades". In reality, there are three units of primary hues. Hue: A hue or "natural" color is the combination of number one colors, where one of the primary colors is at complete depth. The speak of tinting, shading, and tones will make this idea more clean.

5.4 RGB Color Wheel

The RGB model's method to colors is essential due to the truth: One of the thrilling factors of the thought's interpretation of coloration is that it sees the relationship among sunglasses in a round way. A shade wheel is a device that offers a visual illustration of the relationships among all viable shades. The primary colors are prepared round a circle at equal (one hundred twenty degree) periods. (Warning:

Fig. 3 RGB segmentation



Color wheels regularly depict "painter's colors" primary colors, which reason a first rate set of colors than additive coloration (Fig. 3).

5.5 Hue Saturation Value (HSV)

This instantly mirrors the houses of the true shade (true shade) which displays HSV. It defines colors in terms of hue, intensity, and quality (brightness). The usages of hue are:

- The courting among tones across the shade circle is easily recognized.
- Saturation at once correlates to the definition of color in the colors basics.

Section, despite the fact that all saturation does not contain any color, whereas 0 saturation generates white, gray, or black. Depth at this moment represents the notion of deepness in the colors basics level. The desired color position is hues and saturation control (to convert RGB colors mapping to HSV colors or to govern color quantity). Cmap = rgb2hsv (M) responsible for changing an RGB color map M into HSV color map. Color maps are usually design of three matrixes. The element of each color map is within the variety 0–1. The column part of matrix M serves us strength of red, inexperienced green, and blue. Output matrixes Cmap indicate hue, saturation value; therefore, photograph = rgb2hsv (rgb image) converts the RGB photograph same to HSV photograph. RGB is a visual sequence of m-by-three whose three planes contain red, green, and blue. HSV is looking to a path video series, with a hue, a contrast, and photographic speed elements on three planes (Fig. 4).

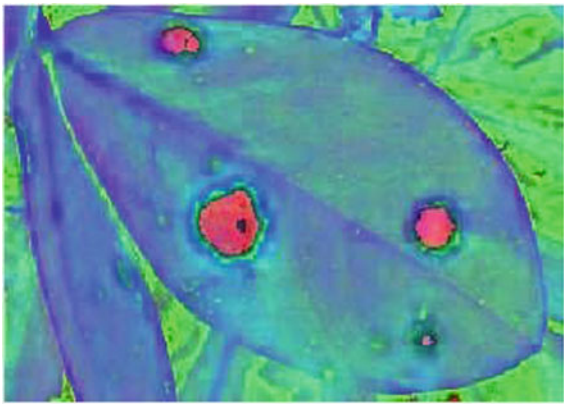


Fig. 4 HSV segmentation

5.6 BPNN and PatternNet

BPNN uses the backpropagation algorithm. It helps in identifying the underlying disease present in the plant leaf. The classifier used is the PatternNet classifier which is a deep neural network classifier. PatternNet takes and returns the pattern reconnaissance neural network (hidden size).

5.7 Input and Output Model

A leaf image is taken as the input. The input goes through the preprocessing phase. The region of interest is identified by converting the input image from RGB to HSV and is isolated by HSV segmentation. The next phase in the process is identifying the name of the underlying disease. This is done using neural networks. A set of images including the ones with disease as well as the healthy ones is the dataset being the used. The patterns detected in the input image (leaf) are matched with the images in the dataset. The input image is classified using the PatternNet classifier (Fig. 5).

The results of the leaf are finally transformed into graphical visualizations through that, and user/researcher can understand from where, how, why the disease is entered. So, that the user/researcher can predict the cause of the disease and take preventive measures to solve the issue before it is too late. The output is the name of the disease (Figs. 6 and 7).

6 Conclusion and Future Enhancement

Plant disease identification model helps in classification of the plant leaf as infected or healthy. The model then further identifies the name of the disease in the leaf. This is

Fig. 5 Damaged leaf

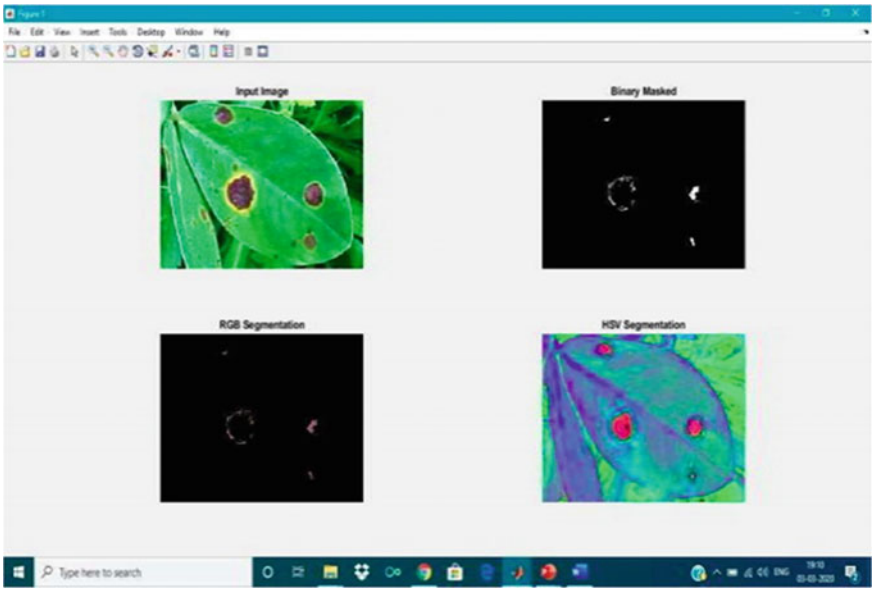


Fig. 6 Attainment of segmentation

preceded by finding the location of the infected region in the leaf through the feature extraction techniques. This model can prove beneficial for farmers in identifying the underlying disease and providing the plants with the appropriate treatment to prevent disease from spreading.

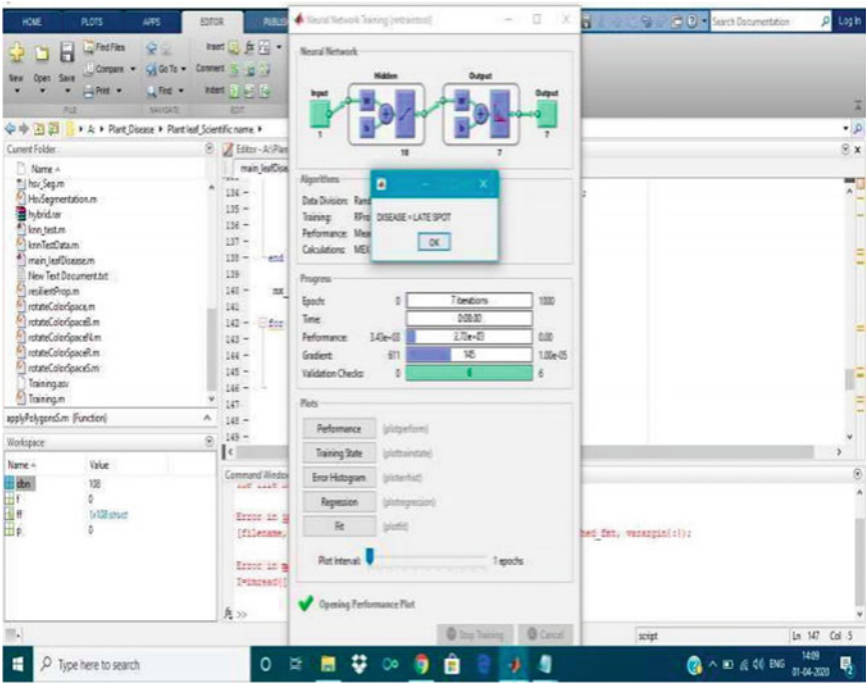


Fig. 7 Detection of diseases occurred

In the future enhancement, it can be done with the help of the IoT or deep learning. In the proposed work, only limited samples are taken, but it can extend to heterogeneous leafs. Also, for predicting the yield of the particular crop or for the particular plants, also, the diseases can be taken into the account.

References

1. Dhingra G, Kumar V, Joshi HD (2018) Study of digital image processing techniques for leaf disease detection and classification. Multimedia Tools Appl 77(15)
2. Shanmugam L, Adline ALA, Aishwarya N, Krithika G (2017) Disease detection in crops using remote sensing images. IEEE Technol Innov ICT Agric Rural Dev (TIAR). https://doi.org/10.1109/tiar.2017.8273696
3. Vijay S, Misra AK (2017) Detection of plant leaf diseases using image segmentation and soft computing techniques. Inform Process Agric 4:141–149
4. Mondal D, Chakraborty A, Kole DK, Majumder DD (20251) Detection and classification technique of Yellow Vein Mosaic Virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier. In: International conference on soft computing techniques and implementations (ICSCTI).https://doi.org/10.1109/icscti.2015.7489626
5. Danping W, Botao W, Yue Y (2013) The identification of powdery mildew spores image based on the integration of intelligent spore image sequence capture device. In: 2013 ninth international

conference on intelligent information hiding and multimedia signal processing. https://doi.org/10.1109/iih-msp.2013.53
6. Dhaygude SB, Kumbhar NP (2013) Agricultural plant leaf diseases detection image processing. IJA-REEIE 2(1)
7. Patil JK, Kumar R (2011) Advances in image processing for detection of plant diseases. J Adv Bioinform Appl Res 2(2):135–141. ISSN 0976-2604
8. Phadikar S, Sil J (2008) Rice disease identification using pattern recognition techniques. In: 11th international conference on computer and information technology. https://doi.org/10.1109/icciteчn.2008.4803079

Impact of Scheduling Algorithms of Wireless Network Control System

Almas Begum and D. Hemalatha

Abstract Network control system (NCS) has become one of the most extensive areas of research in recent years. There are wired and wireless network control system used. Among these, wireless network control system (WNCS) plays a major role in many applications. To make best out of limitations in NCS, WNCS is used to reduce the cost and implementation challenges. This paper focuses on the impact of various scheduling algorithms based on WNCS starting with NCS, evolution of WNCS and its application and also focuses on review of various design techniques used in WNCS including design and scheduling algorithms. Finally, we conclude by identifying the limitations and future enhancements.

Keywords WNCS · NCS · Design techniques · IEEE standards

1 Introduction

The research that is in progress on remote networking anticipates computing to control reforming and remodeling in a way system to control and collaborate along techniques such as IoT, cyber-physical system (CPS), Tactile Internet. Wireless network control system (WNCS) plays a major role stronger compared to the above said techniques. To survive against the limitations of network cables and threads signal fading, mobility, data rate enhancements, minimizing size and cost, user security and quality of service (QoS), WNCS was evolved as it is simpler in design and easy preservation.

WNCS requires innovative model for addressing in association with control and remote frameworks for accomplishing the complete performance and effectiveness of the system, in general, framework execution and productivity. Traditional frameworks of control system configuration depend on the hypothesis to prompt conveyance of sensory information and manage orders that accompany incredibly high reliabilities. The utilization of wireless networks to transmit data recommends

A. Begum (✉) · D. Hemalatha
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_31

Table 1 Features of wireless technologies

Feature	IEEE 802.11	IEEE 802.15.4
Range	10 m	Up to 100 m
Data transmission rate	125 kbps	30 mbps
Devices required to make up network	5–7	Up to 40
Data transmission	Yes can be transmitted	Yes can be transmitted
Error corrections	Forward corrections not possible	Forward corrections not possible
Power utilization	Low utilization	Medium utilization

non-zero deferral and feasibility of error messages. Transmission deficiency or cut-off time missing shall bring outcome depravity in accomplishment of control systems. Therefore, control system's framework needs to incorporate the structure to endure the loss of message problems and also to consider rigid interruption occurring and authenticity issues [1]. The data transmissions must to be adequately dependable and determining the inactivity for the request based on shortest time interval, contingent for duration and limitations of the framework [2]. In addition, eliminating links during data correspondence of detectors and selectors prompts elimination of capacity to the nodes to gain complete adaptability. There are additional limitations that are added-on based on energy and batteries [3].

There are standards defined by IEEE that define the operations of wireless personal area networks called as WPAN in IEEE 802.15.4 and 802.11 called as wireless LAN (WLAN). Table 1 illustrates some of the features of these standards. This standard is used to determine basic lower level network layers that concentrate on minimal effort, low-speed gadgets that communicate everywhere. This is also a standard on which modern, economic and residentiary applications were established being QoS as secondary. The control on cost of the protocol is directly proportional to the probability in loss of message, delay in message and increase in sampling period. The IEEE 802.15.4 convention is a low rate, less cost and less power transmission standard for remote interconnection of fixed as well as compact movable gadgets [4]. There are two types of nodes used in IEEE 802.15.4 standard such as full-function device (FFD) and reduced-function device (RFD) [5].

The aim of the survey is to reveal the requirements and to identify the demands that are correlated with wireless network frameworks of WNCS according to algorithms and protocols to increase the way for effective development. The structure of this survey paper is as follows.

Section 2—WNCS and its applications
Section 3—Design techniques and scheduling algorithm for wireless control system
Section 4—Conclusion.

2 WNCS and Its Applications

A spatial scattered system is in which exchange of data between sensors, actuators and regulators happens through a shared band limited in a communication network. NCS is a combination of control and communication theories where control theory examines about ideal channels that are connected to dynamic systems and communication theory examines about transmission of data over incomplete channels. NCS can be traditionally defined as a conventional feedback loop that is closed in a communication channel and can be shared by different nodes that may be outside the control system in real-world environment.

NCS plays a major role in research over decades and also provides analysis of network structure necessary to implement a reliable, secured communication channel with suitable bandwidth and improvement of data communication protocols for control systems. There are few types of NCS such as: (i) CAN—communication standard framework for increase in data incompleteness and rate at which data is transferred, (ii) Ethernet which is a secondary component with less-cost network that is influenced with the growth of NCS and (iii) WCNS emerged by the need of mobile procedures, adaptable establishments and quick distribution of various applications. Most of control frameworks accept that the information gathered is precise, convenient and lossless [6]. As a result of this mobile procedures, adaptable establishments and quick distribution at the same time decreasing the support cost enhance WCNS to become predominant. The most important characteristics of NCS are to provide an appropriate framework for acquiring the data in real-world environment. Figure 1 illustrates design of WCNS.

For most of the applications, the majority of the congestion occurs on remote networks comprising of certain sensory information from sensory nodes toward at least on an individual controller. So, one of the two resides on the spine is attainable through many spine spots. Thus, as a result of data that passes between sensor nodes and controllers is not really symmetric in WNCS. In addition, various sensors combined to one plant may individually communicate their estimations to the controller [7].

A plant is a constant tangible framework to be controlled. The sources of intake and yields of the tangible framework are constant with timely indicators. Plants yields that exist will be examined by fixed intervals and non-fixed duration intervals by the remote detectors. Every bundle of data related to the framework is communicated to the detector that accompanies a remote network. At this point, detector achieves estimations and processes the control instructions. These instructions that are at that point carried off to the detector appended to the framework. Thus, the circle circuit framework consists both persistent duration and examined elements. Figure 2 illustrates WNCS plants that are in closed loop controlled remotely over a network.

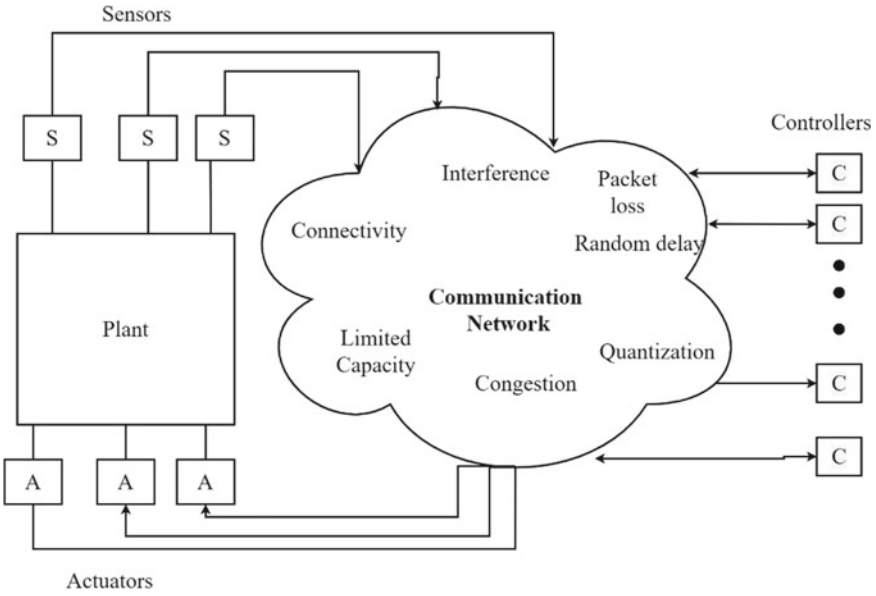


Fig. 1 Design of WCNS

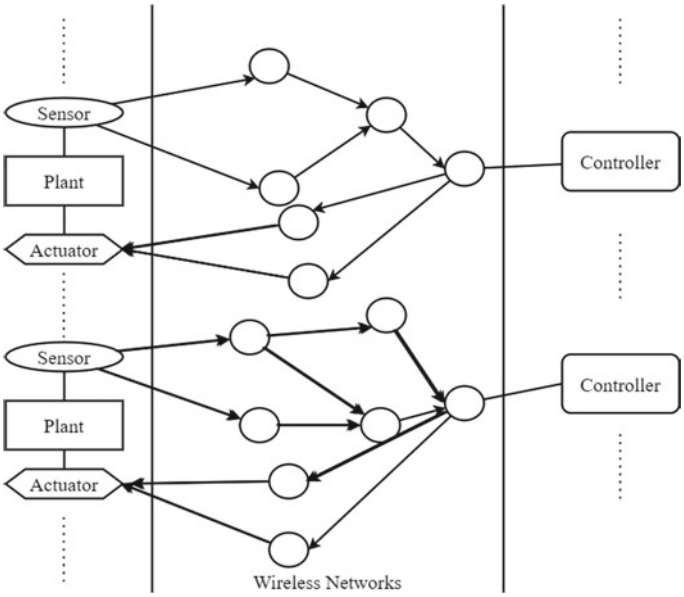


Fig. 2 WNCS plants in closed loop

2.1 Applications of WNCS

There are various types of WNCS applications that are outstanding in evolution of wireless sensor networks with finite data transmission in computing and anticipate to control the applications. Applications of WNCS get motivated in various fields such as vehicular, avionics, building and industrial mechanisms. Intravehicular wireless sensor networks (IVWSNs) consist of components such as control unit, a battery, electronic control units, remote sensors and remote actuators. Wireless sensors are used to transmit data to the analogous control unit as long as it gets energy from nearby control units. Actuators are also used to retrieve the instructions from analogous control units with power on energy-searching gadgets. IVWSN explicit remote sensor network between the vehicle sensors and their relating electronic control unit (ECU) conveyed with the motivation behind either dispensing with the at present existing wires or empowering new sensor innovations that cannot be incorporated into the vehicle utilizing wired methods. Even though there are various techniques explored and determined, the most appropriate and accurate vehicle control system will be energy saving with necessity of sensor nodes at less cost and the shortest distance [8]. Among various vehicular sensor network applications, tire pressure monitoring system (TPMS) is the first application, established on remote communication to strengthen information against vehicle body sensors [9].

WNCS also plays a role in avionics applications that has huge possibility of increasing the performance of aircrafts through cost-efficient operations of flight and reducing the heaviness, cost thereby increasing safety. As there is increased requirement for safety and adaptability, modernized aircraft depends on a huge number of sensors connected and actuators with large number of devices to it [10]. WNCS also plays vital role in building automation and industrial mechanisms. Building mechanism intents to produce ideal-level inhabitant comfort while reducing the energy management. It also requires various range of anticipating facilities to manage temperature, humidity, pressure, etc. In industrialized mechanism, WSAN is used for efficient framework of process controlling in industries, which is a materialized trend followed. Moreover, in the future, various industrial mechanisms require networks based on batteries with "*n*" number of sensors and actuators that broadcast data according to access points.

3 Design Techniques and Scheduling Algorithm for Wireless Control System

There are different techniques for design and optimization of WNCS. The construction of control system is twofold such as one that addresses framework to enhance the network system plant and other is framework of a network system plant with control systems [11]. Basically, remote system model, methods for control systems are broadly classified as interactive design approach and joint design approach [11,

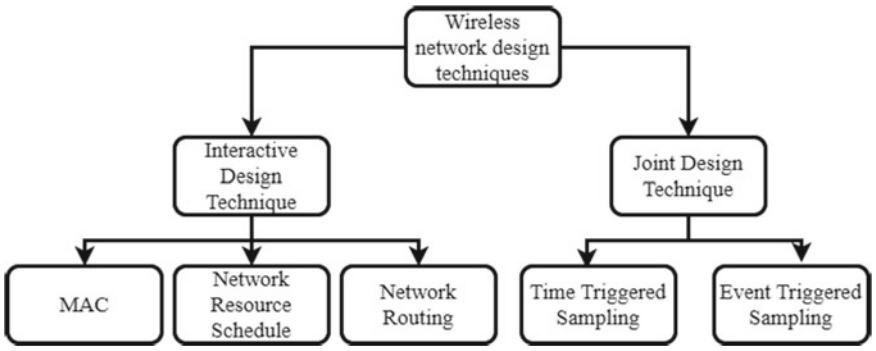


Fig. 3 Techniques used in wireless network design

12]. In interactive approach, the remote networks guidelines are accommodated to fulfill the given collective imperatives required by the control systems for execution. In joint design approach, the remote networks and control system are mutually enhanced. Figure 3 illustrates about the techniques used in wireless design techniques for control system.

3.1 Interactive Design Techniques

Interactive design techniques focus on MAC with network resource scheduling and routing layer, with definite power considering physical layer framework. MAC protocols are classified into contention-based access, schedule-based access and physical layer extension protocols. Standards such as 802.11 and 802.15.4 are used. Contention-based access protocols used in WNCS intent to measure criteria according to CSMA/CA structure based on IEEE 802.15.4. Adaptive algorithm is suggested to reduce the power usage and is implemented in joint design technique. Existing QoS defined in 802.11 of MAC does not consider the cut-off time necessities externally leading to non-predictable efficiency of NCS.

In schedule-based access protocols, it prevents from collisions by observing the schedule broadcast with rigid time occurrences. Sensor nodes are allotted according to time using time-division multiple access (TDMA) and frequency-division multiple access (FDMA) protocols [13]. Time slots are arranged in frames that are repetitive formally. It allows external scheduling transmission with acceptable delay in nodes according to the prioritization. IEEE 802.11 mechanism is used to support domestic application for transmitting and retransmitting of data and also bandwidth to increase the performance efficiently. Physical layer extension proposes scheduling algorithm according to the assigned priority and transmission period that support maximum level of adjustment containing information about loss of data packets in time-triggered and event-triggered samplings.

There are various scheduling algorithms such as dedicated scheduling algorithm, shared scheduling algorithm, priority scheduling algorithm that suggest accurately to designate time slots in order to experience the delay dependability [14]. These algorithms target on packets produced with in the given sampling time and priority that in turn reduce the occurrences of delay in transmission of data packets.

There has been expanding enthusiasm for creating effective multipath directing that advances the system dependability and energy effectiveness of remote systems [15]. Network routing is classified into four designs such as disjoint path routing, graph routing, controlled flooding and energy/QoS-aware routing according to the routing standards. Node-disjoint and link-disjoint are the types of disjoint path routing [16]. Numerous routing path algorithms are proposed that use every node to produce a subgraph that contains shortest path from source to destination [17].

3.2 Joint Design Techniques

The remote system criteria are collectively enhanced seeing the establishment among the performances. They involve examining duration for time-triggered control and level crossings for event-triggered control in the control system. They also include communication competence in physical layer and approaches using MAC algorithm [12]. Joint design technique is categorized in to time-triggered sampling and event-triggered sampling. Time-triggered samplings are based on interactive design techniques. Event-triggered samplings are based on MAC layer [18]. There are some event-triggered sampling that employs CSMA protocol for allotting network resources. Thus, joint design techniques help in assuring interaction to increase the firmness, flexibility and fault tolerance for WNCS.

Thus, there are various commands and instructions that are used to explore how WNCS can be used in control and communications. Though it is not easy to predict the implications, various design techniques can be focused to determine the behavior of WNCS surroundings.

3.3 Scheduling Algorithm

Table 2 shows comparison of different algorithms used for scheduling in WCNS with parameters based on sampling time, priority and delay in transmission. In dedicated scheduling algorithm, packet transmission is based on expected time slots. In shared scheduling, sensors with camera are used for scheduling for controlling the communication channel with closed loops and delay in transmission with 0.02 s.

The sampling period of scheduling algorithm can be evaluated to identify the data

Table 2 Comparison of scheduling algorithm

Algorithm	Priority	Delay in transmission (μs)	Packet transmission
Dedicated scheduling	Low	23.22	Expected time slots
Shared scheduling	Medium	22.67	Share expected time slots
Priority scheduling	High	25.88	Fixed/dynamic

transmission, acknowledgment of data frames and frame space as in Eq. 1.

$$\begin{aligned} \text{Tt} \geq \text{backoff duration} + \text{data frame} + \text{ack for time} \\ + \text{frame duration} + \text{inter frame space} \end{aligned} \tag{1}$$

Similarly, delay in transmission is based on closed-loop delay that is evaluated with components such as sensors, connectors and actuators as shown in Eq. 2.

$$\tau = \tau\text{sensor\_ connector} + \tau\text{connectors} + \tau\text{connector\_ actuators} \tag{2}$$

3.4 Observations

There are various standards that have been implemented for wireless control network systems. Although there are relevant advantages, there are some characteristics that provide less rate at which data is transmitted, less power resources utilized and still experiences multiple challenges. Using IEEE 802.15.4, arbitrary waiting time for packet transmission can be delayed using collision avoidance methods. Collision can be avoided by either using closed-loop systems or by other means of applications. Based on the priority of the packet transmission, interruptions can be avoided.

4 Conclusion

WNCS serves as basic technology that concentrates on safety of the control systems in various applications such as vehicular, avionics, building and industrial mechanisms. This study paper presents aspects of NCS and WNCS with some of the protocols and design techniques used. A brief review on design techniques is used in WNCS. Thus, it enables more productivity, expandability and less cost. Simulation techniques can be used to achieve the inferences about the failures that may influence behavior of

WNCS. To carry forward with future directions, different independent components such as sensors, actuators can be used to improve the efficiency of network and also to enhance the energy resources.

References

1. Willig A (2008) Recent and emerging topics in wireless industrial communication. IEEE Trans Industr Inf 4(2):102–124
2. Gungor V, Hancke G (2009) Industrial wireless sensor networks: challenges, design principles, and technical approaches. IEEE Trans Industr Electron 56(10):4258–4265
3. Hong Y (1998) Networked control systems [Online document]. Available http://www.enme.umd.edu/icelab/ncs/ncs.html
4. IEEE 802.15.4-2006 (2006) Part 15.4: wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs)
5. Xia F, Vinel A, Gao R, Wang L, Qiu T (2011) Evaluating IEEE 802.15.4 for cyber-physical systems. EURASIP J Wirel Commun Network 2011:1–14
6. Millan YA, Vargas F, Molano F, Mojica E (2011) Wireless networked control systems review. In: Electronic and telecommunications engineering program, Oct 2011, pp 1–7
7. Matveev AS, Savkin AV (2003) The problem of state estimation via asynchronous communication channels with irregular transmission times. IEEE Trans Autom Control 48(4):670–676
8. Demir U, Ergen SC (2016) ARIMA-based time variation model for beneath the chassis UWB channel. EURASIP J Wirel Commun Network 2016:178
9. Velupillai S, Guvenc L (2007) Tire pressure monitoring [applications of control]. IEEE Control Syst 27:22–25
10. Technical characteristics and spectrum requirements of wireless avionics intra-communications systems to support their safe operation. ITUR, Technical report 2013, m.2283-0
11. Al-Dabbagh AW, Chen T, Design considerations for wireless networked control systems. IEEE Trans Ind Electron. https://doi.org/10.1109/TIE.2016.2564950
12. Park P, Ergen SC, Fischione C, Lu C, Johansson KH, Wireless network design for control systems: a survey. IEEE Commun Surv Tutorials. https://doi.org/10.1109/COMST.2017.2780114
13. Segun A, Akinwunmi AO, Ogunti EO (2015) A survey of medium access control protocols in wireless sensor network. Int J Comput Appl 116(22):0975–8887
14. Seno L, Cena G, Scanzio S, Valenzano A, Zunino C (2017) Enhancing communication determinism in Wi-Fi networks for soft real-time industrial applications. IEEE Trans Industr Inf 13(2):866–876
15. Ploennigs J, Vasyutynskyy V, Kabitzsch K (2010) Comparative study of energy-efficient sampling approaches for wireless control networks. IEEE Trans Industr Inf 6(3):416–424
16. Ganesan D, Govindan R, Shenker S, Estrin D (2001) Highly-resilient, energy-efficient multipath routing in wireless sensor networks. ACM SIGMOBILE Mob Comput Commun Rev 5(4):11–25
17. Jindong Z, Zhenjun L, Yaopei Z (2009) ELHFR: a graph routing in industrial wireless mesh network. In: IEEE ICIA
18. Tian G, Camtepe S, Tian YC (2016) A deadline-constrained 802.11 MAC protocol with QoS differentiation for soft real-time control. IEEE Trans Industr Inf 12(2):544–554

Scientific Exploration of Conceptual and Algorithmic Terminologies of Machine Learning

Juginder Pal Singh

Abstract At beginning, machine learning (ML) was only the endeavor of some computer geeks who wanted a program that can learn to play games by itself and the field of statistics which never employed any approaches based on computations. Now, it has developed into an independent field of study that has provided us base to develop self-learning functions. ML has also developed various algorithms for interpretation of text, email filtering, recognizing pattern, and several different commercial aspects. It also led to a different field of research such as the field of data mining which identifies hidden and non-hidden regularities and irregularities in so fast-growing data and artificial intelligence which advanced machine learning algorithm like neural networks to create a program that can closely mimic human beings. The research proposed here throws light on explaining various conceptual terminologies and expansion of machine learning from where it has started; among the major popular machine learning algorithms, some are discussed here that are being used today and for future scope of machine learning.

Keywords Learning · Human learning · Machine learning · SVM · Cognitive services

1 Introduction

One of the definitions of "learning" is to study past experience to improve future performance. In machine learning, past data is analyzed to give the output, and the accuracy of it improves as more data is inputted. The primary goal in ML is to create automatic learning functions. ML algorithm builds a mathematical model based on data which is inputted known as "training set". This training set is then used for automatic prediction and decision-making (Fig. 1).

J. P. Singh (✉)
Department of Computer Engineering & Applications, GLA University, Mathura 281406, India
e-mail: juginder.singh@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_32

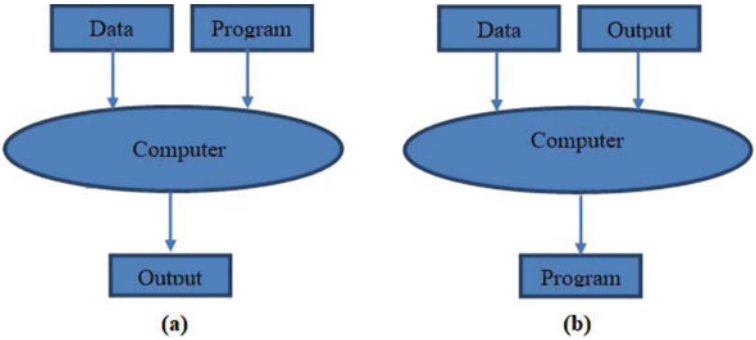


Fig. 1 **a** Traditional programming **b** Machine learning

How this training set is used depends on what ML algorithm you are using. Today, ML is used in various applications like Google search and email filtering where it was almost impossible to create conventional algorithms [1, 2].

A. **Machine Learning**: Birth of machine learning happened when computer science and statistics were joined together. Creation of machine that is able to solve a particular problem is the domain of computer science, while modeling and measuring reliability of data is domain of statistics. The idea of machine learning takes inspiration from these two domains to create a machine that can automatically work on modeled data. In computer science, we have to program the whole system independently to do a task, while in ML, we have to create just the learning algorithm, and rest is done by the algorithm itself. Machine learning augments on statistics domain of inference and probability of data and includes additional features like feasibility of algorithms for data processing and compoundation of tasks and increased performance count [3, 4] (Fig. 2).

B. **Machine Learning and Natural (Human) Learning**: Several researches show that there is a close relation between machine learning and human learning in various interdisciplinary field on inter-sciences. Machine learning algorithm learns in a same way life that an in fact child learns, i.e., through trials and errors. These researches have yet prove significant as algorithm design on human learning are much weaker than statistics-based ML algorithm. The reason can

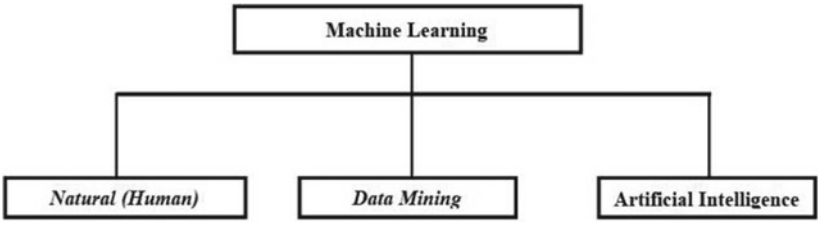


Fig. 2 Emerging areas of ML

be traced to fact that we still long way away from uncovering the mystery of human learning. Although, recently some strides has been taken in advancement in these fields [5, 6].

C. **Data Mining and Machine Learning**: With development of data mining, the relationship between the two has become so interrelated that it is hard to tell when data mining is used or when machine learning is being implemented. Data mining is a cross disciplinary field that excels in searching the properties of a given data. It interprets a data and analyzes the correlation between data and also recognizes pattern in the data. To summarize, data mining only focuses on discovery of data not prediction which is the focus of machine learning. Data mining can use different ML algorithms to do mining. Similarly, in many cases, it is also used to preprocess data for ML algorithm to be implemented [7].

D. **AI and Machine Learning**: As already mentioned in this paper, machine learning was developed when some computer enthusiasts wanted to achieve "artificial intelligence". The fact that "ML is a subset of AI" is not accepted by many researchers and scholars. ML first used symbolic methods called "neural networks" to first create artificial intelligence. But, as human started realizing other application of ML, the rift between the two came. AI still aims in developing algorithm for artificial intelligence, whereas ML starts using more probabilistic approach [8].

2 Present Research Demands of ML

Numerous researches have been already done on machine learning in the past decade. Following are some of scopes of improvement for machine learning where research is still being done:

A. **Use of Unlabeled Data in Supervised Learning**. Supervised learning is an algorithm that maps an input to an output with knowledge of previous input–output pairs. This mapping is done with help of labels present in data, so supervised learning is mostly performed on labeled data. But practically, there is very less labeled data in the world. Most of labeled data are building with help of preprocessing and labeling the unlabeled data which incurs huge cost. One way to reduce this cost is to make a supervised algorithm that works on unlabeled data. According to research done by many institutions, there are several cases like spam filtering where having unlabeled data for supervised learning is actually more beneficial for the algorithm. So, supervised ML algorithms are being developed that can handle unlabeled data [9].

B. **Transferring the Learning Experience**. Real-life problems are not simple and short. ML algorithm had to be applied on large sets of data. For example: If a company wants to collect data on various diagnostic functions all over India, the diagnostic function built on various patient treated by the hospital of Lucknow will be similar but not necessarily same as Delhi. So, making of diagnostic function that applies to whole nation is very difficult. Few methods

are used to find the similarity between two diagnostic functions. For example: An approach given by hierarchical Bayesian assumes that learning parameters of both Delhi and Lucknow have some common probabilities [10].

C. **Linking Different ML Algorithms**. There is large number of ML algorithms used in various fields all over the world. Research has been done to find correlations between existing ML algorithms. For example: Naive Bayes and logistic regression, these two supervised algorithms approach data distinctly. But, for a specific type of training data (i.e., when the criteria of naive Bayes classifier are fulfilled and the number of examples in trying set tends to infinity), their equivalence can be seen. The conceptual understanding of ML algorithms, their convergence features, their effectiveness and limitations, and their applications are the fields of research that need to be further studied and experimented.

D. **Machine Learning Algorithm Which Collects Their Own Data**. However, good a ML is, it will not be fully automated and revolutionary if we have to feed data in it every time manually. Several research is being done for making a ML which collects its own data for processing and learning. Many effective strategies are analyzed which hand over the task of collecting data to ML algorithm. For example: an algorithm which is able to learn the success of vaccine by monitoring the infected patients (let us assume patient suffering from coronavirus) and vaccine injected into them. Algorithm will also analyze any unknown side effects of the vaccine and will try to minimize it [11, 12].

E. **Privacy Preserving Data Mining**. As the number of cybercrime is increasing day the day, the preservation of privacy in data mining has become very important. Thus, the need to develop an algorithm that is able to mine the data without exploiting the underlying information is in demand. One of such algorithm is privacy preserving data mining (PPDM) which protects the data by changing them by masking the original vulnerable data so that it cannot be exploited. Then, it uses various methods to determine the original data up to a certain level of accuracy. It is highly effective algorithm for privacy. The only tradeoff is that it leads to loss of accuracy of data little bit, but by using large amount of data, it can be minimized (Fig. 3).

3 Category Divergence of ML Algorithms

A. **Learning Style Category**

1. **Supervised Learning**. As indicated by the "supervised" word in the name, this ML algorithm takes place in the presence of a supervisor. This supervisor is the labeled set of data which helps in building of classifier which is then properly trained to determine the label of test data. Basically, a supervised learning is an algorithm that maps an input to an output with knowledge of previous input–output pairs. Classification and regression are the most popular uses of supervised machine learning.

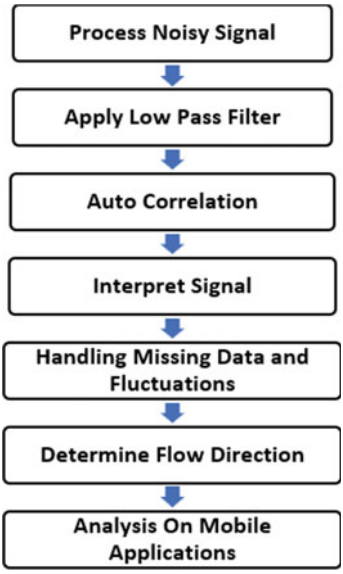


Fig. 3 Flow of control to correlate and process noisy signal

2. **Unsupervised Learning**. No guidance is given in this ML Algorithm, i.e., the data provided as input is not a labeled data. Algorithm sorts data with help of differences and pattern without having knowledge of any classifier built on prior data. Clustering and association are the most popular uses of unsupervised machine learning [2] (Fig. 4).
3. **Semi-supervised Learning**. This machine learning algorithm is mash up of both supervised and unsupervised ML algorithms. Both labeled and

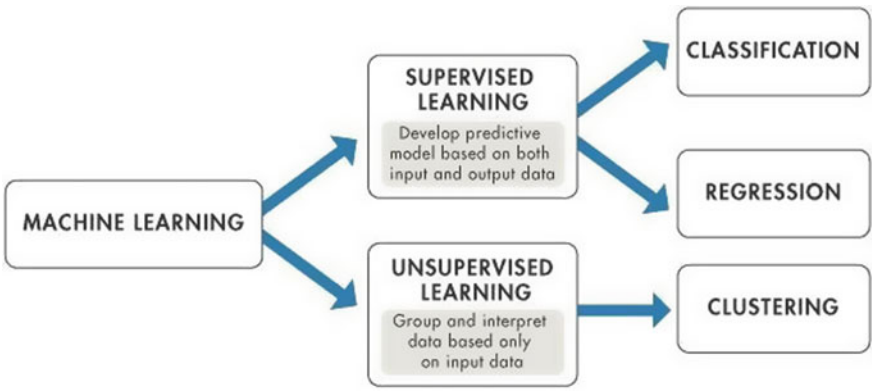


Fig. 4 Classification of machine learning

unlabeled data are used, and pattern finding is also done in the data. It counters the disadvantage of both supervised, i.e., dataset being hand labeled and unsupervised, i.e., limited use [4].

4. **Reinforcement Learning**. This machine learning algorithm uses the concept of "maximization of reward" and "learning by experience". Here, there is no correct answers given to machine to train the model, instead the reinforcement coordinator decides whether the answer is right, i.e., closest to the correct answer, the output is then fed to reinforce the algorithm known as maximization of reward.

B. **Similar Algorithm Category**

1. **Regression Algorithms**. The co-relation between output values and variables is exploited by this algorithm. The variables are mostly independent, but output had to have some kind of dependency on the variables. This algorithm using this co-relation builds a model and uses it to predict values for new inputs. Some Ml algorithms which use this regression are: linear regression, ridge regression, polynomial regression, logistic regression, etc.
2. **Instance-Based Algorithms**. It is also known as memory-based algorithm as it stores variables and the result of all the data input in the algorithm. This variable-result pair is known as instances. To predict the result of input in case of new data, these instances are analyzed. The algorithm works on principle "winner takes everything", i.e., instance can be replaced, which other instances which are the best fit. ML algorithm which uses this instance-based algorithm is: learning vector quantization, *K*-nearest neighbor (popularly known as KNN), etc.
3. **Regularization Algorithm**. Rather than being a complete ML algorithm, this technique or algorithm is a popular aid which is used along with other ML algorithm to reduce overheating, i.e., ML starts learning from inaccurate data entries in the dataset. The regularization algorithm reduces the degree of polynomial function in another Ml algorithm which is a root cause of overeating. There are many ways to reduce the degree of polynomial, but the best way is by making the coefficient of the highest order terms in a polynomial smaller. Examples of regularization algorithm are: lasso regression, ridge regression, etc. [8].
4. **Decision Tree Algorithms**. It is a type of supervised learning algorithm which uses a tree representation to categorize data in which each node acts as a class label on which decision is made. Many internal node branches of this node are representing distinct attribute set according to the class label. These steps are continued recursively until a prediction is made. In this process, it forms a tree-like structure which is the result of decisions taken at every node, hence, named "decision tree". It is faster than any other algorithm for classification. A major decision tree algorithm is classification and regression tree (CART) [9].

5. **Bayesian Algorithms**. This is one of the most basic algorithms of machine learning. It uses Bayes algorithm to predict the result. It is a probability-based algorithm that assumes all the variables are independent. Examples of Bayesian algorithms are: naive Bayes, Bayesian network (BN), Gaussian naive Bayes, etc.
6. **Support Vector Machine (SVM)**. It is one of most popular machine learning techniques known for producing highly accurate results with less computing power. In SVM, our main aim is to find a hyperplane. This plane is found in an N-dimensional space where each dimension represents a different variable given as input. The best fit hyperplanes are chosen, and they help to classify the data by acting as a separator. Support vectors (data points nearer to the hyperplane) can heavily influence hyperplanes. Both hyperplane and support vector together build our SVM.
7. **Clustering Algorithms**. It is an unsupervised ML algorithm which uses pattern present in data to group them into different group. Each group is distinct from the other group, and elements of each group exhibit similar properties. ML algorithm which uses this instance-based algorithm is: *K*-means clustering algorithm, mean-shift clustering algorithm, DBSCAN, etc.
8. **Association Rule Learning Algorithms**. It is an algorithm which is used to find an association and relationship in large data. It uses help of certain rules which aid in finding these relationships. These rules are based on one general principle, i.e., to find something "interesting". The term "interesting" refers to unique association pattern in the data. The algorithm is mostly used in market analysis. Examples of association rule learning algorithms is a priori algorithm, Éclat algorithm, etc.
9. **Artificial Neural Network (ANN) Algorithms**. ANN is inspired by the working of human neural networks. ANN is basically a set of algorithms which is capable of recognizing patterns. Patterns are recognized from vectors. Vectors are numerical representations of any form of data. The basic aim of neural network is to cluster and classify. Unlabeled data with similarities present between them can be grouped and classified through ANN. More complex form of neural network called deep learning can even extract features from other ML algorithms. Examples of ANN are: perceptron neural network, backpropagation, radial basis function network (RBFN), etc.

4 Future Scope

A. **Improved Cognitive Services**. While right now with the help of machine learning, we are able to include intelligent capabilities in our devices. These intelligent machines can recognize vision, speech and can even distinguish between different dialects. Apple's Siri and Amazon Alexa are good examples

of it. They are now capable of assisting humans to a certain limit in their daily life. Some even more powerful AI can even converse with human to a certain extent. In the future, these intelligent devices will be so powerful that they can even be able to read the human's mood just by analyzing our speech pattern and converse with us with the same level of intellect as any other human or even more.

B. **Quantum Computing**. Even prominent people like Satya Nadella, CEO of Microsoft, and Elon Musk, CEO of Tesla, have come out and said that quantum computing will reshape the world of computing. Traditional computer uses binary. Even complex problems are broken in form of 0s and 1s. No doubt they are fast, but they all have computational limits. Quantum computing will break these limits as much faster and uses less resources. This allows our ML to evolve and can even become more sophisticated. In the future, quantum computing can help machine learning to resemble a network similar to our neural network.

C. **Robots**. Traditionally, robots were machine designed to perform a certain specific task only. But, with the rise of machine learning, they are becoming highly sophisticated. Robot vision, robot sense, and self-supervised learning are great achievements in the field of robots. With their help, drones are not capable enough to fully automate delivery, many factories have become fully automated and need almost negligible human intervention, and cars like Tesla are now fully automated and capable of self-driving. Soon in the near future, humans will have our own personnel robot assisting them in their day to day life.

5 Conclusion

Primary aim of machine learning authors and application developer along with researchers of this field is to build and design more advanced, proficient, and time- and space-optimized practically achievable all purpose learning mechanisms which could generate best result over a wider domain spread. In machine learning context, the performance realized by a method to efficiently use data resources is also a very important paradigm. Better prediction of accuracy along with prediction rules also plays a vital role in machine learning. Being fully data-driven and the ability of examining a huge count of data in very little span of time, machine learning also proves much better over manual human programming. Also, machine learning proves to be more accurate, error-free, and lesser prone to human bias. Keeping in mind the above scenarios, solution developed to solve tasks that uses sensors such as speech recognition, computer vision can be easily achieved via the concepts and algorithms of machine learning.

References

1. Singh JP, Mishra MK, Khan MA (2017, July) Energy-efficient approach towards video-based sensor networks (wireless) beneath barrier coverage. In: 2017 8th International conference on computing, communication and networking technologies (ICCCNT). IEEE, pp 1–5
2. Mduma N, Kalegele K, Machuve D (2019) A survey of machine learning approaches and techniques for student dropout prediction
3. Vashistha P, Singh JP, Khan MA (2020) A comparative analysis of different violence detection algorithms from videos. In: Advances in data and information sciences. Springer, Singapore, pp 577–589
4. Ameri S, Fard MJ, Chinnam RB, Reddy CK (2016, Oct) Survival analysis based framework for early prediction of student dropouts. In: Proceedings of the 25th ACM international on conference on information and knowledge management, pp 903–912
5. Singh JP, Gupta A (2020, July) Scientific exploration of hand gesture recognition to text. In: 2020 International conference on electronics and sustainable communication systems (ICESC). IEEE, pp 363–367
6. Chapelle O, Sindhwani V, Keerthi SS (2008) Optimization techniques for semi-supervised support vector machines. J Mach Learn Res 9(Feb):203–233
7. Mitchell T, Cohen W, Hruschka E, Talukdar P, Yang B, Betteridge J, Krishnamurthy J (2018) Never-ending learning. Commun ACM 61(5):103–115
8. Cour T, Sapp B, Taskar B (2011) Learning from partial labels. J Mach Learn Res 12:1501–1536
9. Chen X, Shrivastava A, Gupta A (2013) Neil: extracting visual knowledge from web data. In: Proceedings of the IEEE international conference on computer vision, pp 1409–1416
10. Kumar M, Bhatnagar C (2017) Crowd behavior recognition using hybrid tracking model and genetic algorithm enabled neural network. Int J Comput Intell Syst 10(1):234–246
11. Kumar M, Bhatnagar C (2017) Zero-stopping constraint-based hybrid tracking model for dynamic and high-dense crowd videos. Imaging Sci J 65(2):75–86
12. Kumar M, Bhatnagar C (2017) Hybrid tracking model and GSLM based neural network for crowd behavior recognition. J Central South Univ 24(9):2071–2081

Deep Learning in Human Activity Recognition from Videos: A Survey

Neeraj Varshney

Abstract Human activity recognition is a trending area of research in current times. It has variety of applications especially from security point of view. Recognizing human actions from video is a tedious task in computer vision due to largely invariant nature of scenarios that occur and also the infinite amount of data that needs to be worked upon that too accurately and fast. Activities are classified into two major groups. The simple activities come under single-layer group wherein simple activities like sitting, walking, running are performed by one or two subjects. Another group is called hierarchical group that consists of much complex activities or activities in sequence being performed by one or many subjects at a time. This paper discusses some papers associated with human activity recognition from videos. We examined each paper to highlight the method used for feature extraction and classification. We also noted the method used and accuracies achieved by the authors on various public datasets.

Keywords Deep learning · Activity recognition · Video

1 Introduction

Computer vision is a field of study that tends to develop methods and techniques required for the computer to see and understand the world, and the input given is being in the form of images and videos. In these times, images and videos are abundant in number because of cameras being available and used wide scale around the world. Due to this reason, computer vision in recent times has seen tremendous growth. Computer vision can be further said to be a sub-field of artificial intelligence and machine learning. Computer vision has applications across a wide variety of fields such as the automotive industry, health care sector, banking, agricultural sectors.

Human activity recognition is a problem of classifying data obtained from sensors or videos into known well-defined movements performed by humans. Recognizing

N. Varshney (✉)
Department of Computer Engineering and Applications, GLA University, Mathura 281406, India
e-mail: neeraj.varshney@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_33

of human activities is a challenging task because there is no single and precise way or formula to determine specific actions. Moreover, in order to get accurate result, large volumes of data is analyzed at a given moment. Earlier, hand-crafted features were used, but more recently, deep learning models are being developed which give us better accuracy. In deep learning, features are set up automatically; therefore, it is an unsupervised way of dealing with the problem. Human activity recognition can be used to build smart health care centers, secured public places, wildlife observation centers, and many more.

In this paper, we study and summarize the recent trends that are taking place in human activity recognition (HAR) especially from videos. We analyzed each paper to highlight the main advantages that each paper have. Also, we noted the input data, feature detector, classifier, and the accuracies achieved by each paper on respective public datasets.

2 Literature Review

Human activity recognition (HAR) is an essential technology for developing relevant applications such as visual surveillance and human computer interaction. Usually, atomic actions are recognized in RGB videos, but with the availability of RGBD data, Lillo et al. [1] introduced a hierarchical compositional model that operates at three levels of abstraction. Geometric and motion descriptors are at lowest level, sparse compositions at intermediate level, and spatial and temporal compositions at highest level for representing complex human activities. However, limitation of this work is that annotated data is required at the level of action which can be an issue for an application involving large volumes of data. This issue of limited annotated data is addressed by Sarker et al. [2] in which a novel framework is proposed that performs HAR from RGB video having limited labeled data. A similar hierarchical spatiotemporal has also been discussed by Xu et al. [3]. Video-based security subtractions could be more efficient if the background noise is reduced and only the subjects are kept into focus. This method of background subtraction is presented by Zaidi et al. [4]. The method is able to detect objects from a recorded video. However, it has not been applied on real-time live video. Therefore, Samir et al. [5] present a RGB input system for recognizing different real human activities using a stationary camera. Fine-grained action recognition is much harder to detect than general human actions because they occupy relatively small-scale image regions. Ma et al. [6] seek to improve fine-grained action discrimination along with general action recognition. Advantage of the method is that it is able to reduce feature dimension at the same time combining appearance and motion information.

Skeleton and joint-based human activity recognition is discussed by Nikolov et al. [7] and Qi et al. [8]. Both the works used RGBD videos as inputs. Advantage of using the above discussed methods is that complex features with spatiotemporal configurations of subjects in video are automatically learned. Phyo et al. [9] used deep learning for human activity recognition using motion of joints. The system is

promising as far as real-time experimentation is considered due to fast processing time.

Jalal et al. [10] used body silhouettes and individual skeletal points to develop a novel framework providing robust multifeature descriptors for activity recognition. However, complex scenarios are not considered while experimentation. A more robust multimodal feature-level fusion method is introduced by Ehatisham-ul-Haq et al. [11] where feature-level fusion of time-domain features and HOG features reduces the computational complexity and improves the recognition. However, the method fails to incorporate multiview activity recognition. Convolutional neural networks are widely used in human activity recognition especially in classification process. Zhou et al. [12] and Zalluhoglu and Ikizler [13] used CNN. However, the models lack explicit modeling of sequential information other than 3D temporal convolutions. This issue is addressed by Long et al. [14] where a model called asymmetric residual neural network is discussed. It includes recurrent neural networks and LSTM memory networks. RNNs can find patterns with long-term dependencies because it can memorize the information extracted from past data. Binary quantization map (BQM) descriptors, R-transform, and Gaussian mixture models are discussed by Gnouma et al. [15], Dhiman and Vishwakarma [16], and Najar et al. [17], respectively.

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2019, Moencks et al. [18]	RGBD	Use a multimodal sensor dataset, context-dependent feature extraction can be done through data preprocessing algorithm	Joint extraction algorithms	Deep neural network	HAR identification through multimodality	LboroH AR	96.8
2019, Long et al. [14]	Sensor data	At the same time, ARN considers both spatial and temporal features of data	CNN	CNN	ARN is a kind stable model which is robust for the lengths of slide time window	Opportunity, UniMiB-SHAR	90.29, 76.39
2018, Najar et al. [17]	3D data	Unsupervised learning methods were used for HAR of finite multivariate Gaussian mixture model	3D SIFT, Harris, *K*-means, PLSA	FP-MGGMM-MML	Unsupervised learning approach for human activity recognition	UIUC sports, human interaction images, KTH	97.96, 89.49, 91.97

(continued)

(continued)

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2020, Gnouma et al. [15]	Video frames	Introduce to predict and localize anomaly events, introduce a new framework for finding motion extraction, i.e., BQM	Sparse autoencoder	Softmax	During the training, prominent region to be identified, using a new BQM descriptor	UCSD, avenue	97.8, 96.2
2019, Zalluhoğlu and Ikizler [13]	Video frames	Collective activity recognition is identified through multistream architecture	CNN	Softmax	Fuse region of interest and CNN streams in early phases are used to produce effective results in recognizing activities	Collective activity, volleyball	88.9, 72.4
2019, Ehatisham-ul-Haq et al. [11]	RGBD + acc + gyro	Fusion at feature level has been proposed to identify HAR. Two different sensing modalities vision and inertial data are used	HOG	KNN	Enhance accuracy of recognition through the fusion of time-domain features and HOG features from RGBD videos	UTD-MHAD	98.3

(continued)

(continued)

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2018, Ma et al. [6]	RGB video frames	A novel idea about region sequence-based six stream CNN feature for HAR in videos is used which combines variant scales of image information and video motion information	CNN	SVM	Dimensionality of the feature reduce also framework is used to combine motion information	Sub-JHMDB, MPII cooking	76.9, 70.3
2016, Lillo et al. [1]	RGBD	A model based on hierarchical compositional was introduced, works at 3 levels of abstraction	*K*-means, HMM	SVM, HMM, H-BoW	Their approach was good to handle occlusion and provide meaningful annotation	MSR Action 3D	92.2
2018, Qi et al. [8]	RGBD	An approach based on sparse representation and dictionary learning is used for self-learning of joint configuration	Sparse encoding	SVM	System self-learned the complex features by the spatiotemporal geometric configurations of body joints	CAD 60	94.12

(continued)

(continued)

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2017, Samir et al. [5]	RGB	Human activities are monitored by a single stationary camera in real time	Fourier descriptor	SVM	System resolves the issue of variation in different body movements of human and the way one performs an activity. Model was able to categorize the suspicious and normal activity	Weizmann	94.46
2018, Sarker et al. [2]	RGB	An innovative idea was proposed for activity recognition that uses skeleton data extracted from RGB video and deep BLSTM model	Open pose API (skeleton)	BLSTM, Softmax	A cost effective solution which can be sued in the ordinary cameras	KTH, UTD-MHAD	96.07, 91

(continued)

(continued)

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2019, Phyo et al. [9]	Depth data	A new idea was given for human activity recognition in the consumer electronics world by the use of color skeleton, MHI, and relative joint image	3D CNN	Softmax	Better utilized on environment condition/domain structures with no restriction and also can be used or real-time application because of the fast processing time	CAD 60, UTKinect Action 3D	96.15, 97
2019, Dhiman and Vishwakarma [16]	RGBD	R-transform and Zernike moments on average energy silhouette images were used to construct a robust feature vector	R-transform + Zernike moments	KNN/SVM	Obtained a complete geometrical transformation invariant feature with less noise compassion	ARA, UR fall detection, KARD, NUCLA	96.5, 96.64, 95.9, 86.4
2019, Jalal et al. [10]	RGB	CNN-based multifeature algorithm designed	Silhouette	CNN	Better for complex realistic environment	BIT interaction	84.63
2016, Xu et al. [3]	RGB	A hierarchical spatiotemporal model is used by modeling spatial and temporal constraints	SVM	HSTM	Obtain superior classification by using the spatial similarity and temporal similarity of activities	BIT, UCF, CASIA, UT, Gupta	88.50, 90.67, 95.24, 94.17, 96.30

(continued)

(continued)

Paper ref. and author	Input	Description	Feature extraction	Classifier	Advantage	Dataset	Accuracy
2019, Zhou et al. [12]	Sensor data	Method is efficient for indoor localization. CNN is used for activity recognition based on smart phone	CNN	CNN	Automatically learns features		
2017, Gowda [19]	RGB	An approach to HAR using a combination of deep belief networks	Weber local descriptors	CNN	Proposed approach used modification of the Weber local descriptors, and deep network is used to extract features from video and frames along with the modification of LBP		
2017, Cherian et al. [20]	RGB	A novel learning-to-rank formulation for capturing the temporal evolution of actions in video sequences by learning subspaces	CNN	CNN	An efficient conjugate gradient optimization scheme for pooling		

3 Comparative Study

We focused on papers using machine learning, deep learning, and feature-based learning in order to recognize human activities from videos. What we have found so far is that all the architectures vary as per the temporal dimension incorporation in the data and also the amount of data available as input for the architecture. Dhiman and Vishwakarma [16] used KNN and SVM separately in order to get brilliant accuracies for an input RGBD data. This means construction of a robust feature vector in this case by R-transform is beneficial even if the classifier used is only simple machine learning algorithm. An architecture considering both spatial and temporal components separately and then combining the end observation to bring out accurate human activity recognition is shown by Xu et al. [3]. It got good results in five different public datasets. Xu et al. [3] use 3D convolutional filter as against Zalluhoğlu and Ikizler [13] and Ma et al. [6] which uses 2D convolutional filters to get far better accuracies on multiple public datasets. In general, when input is in the RGB form only, it can be seen that deep learning models perform better than conventional machine learning models, as seen in Gnouma et al. [15] and Sarker et al. [2] where deep learning models are used as against conventional machine learning architecture used in Ma et al. [6] and Samir et al. [5]. However, what is to be remembered is that in all the cases whatever be the dataset used, the datasets are built under different circumstances. A dataset might use single view, multiview, single background, single or multiple number of subjects under consideration.

4 Conclusion

In this paper, we presented a quantitative and theoretical analysis of various machine learning, deep learning, and feature-based learning methodologies used previously for recognition of human activities from videos. A successful recognition does not only involve the classification method used but also the efficient use of feature detection techniques and correct labeling of sub-activities. Also, background noise is an inescapable element from which the subject of interest is needed to be separated in all the cases. It has been observed that the labeling of sub-activities and activities as a whole is a tedious task that involves much more computation than any other process, and deep learning methodologies are able to perform this task in an unsupervised way. As a result of which, given similar input data, deep learning models being hierarchical in nature brought out more accuracy than conventional machine learning and feature-based learning techniques.

References

1. Lillo I, Niebles JC, Soto A (2016) Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos. Image Vis Comput 59. https://doi.org/10.1016/j.imavis.2016.11.004
2. Sarker K, Masoud M, Belkasim S, Ji S (2018) Towards robust human activity recognition from RGB video stream with limited labeled data
3. Xu W, Miao Z, Zhang X-P, Tian Y (2017) Learning a hierarchical spatio-temporal model for human activity recognition 1607–1611. https://doi.org/10.1109/ICASSP.2017.7952428
4. Zaidi S, Jagadeesh B, Sudheesh KV, Audre AA (2017) Video anomaly detection and classification for human activity recognition. In: 2017 International conference on current trends in computer, electrical, electronics and communication (CTCEEC). https://doi.org/10.1109/ctceec.2017.8455012
5. Samir H, Abdelmunim H, Aly GM (2017) Human activity recognition using shape moments and normalized Fourier descriptors. In: 2017 12th International conference on computer engineering and systems (ICCES). https://doi.org/10.1109/icces.2017.8275332
6. Ma M, Marturi N, Li Y, Leonardis A, Stolkin R (2018) Region-sequence based six-stream CNN features for general and fine-grained human action recognition in videos. Pattern Recogn 76:506–521
7. Nikolov P, Boumbarov O, Manolova A, Tonchev K, Poulkov V (2018) Skeleton-based human activity recognition by spatio-temporal representation and convolutional neural networks with application to cyber physical systems with human in the loop. In: 2018 41st International conference on telecommunications and signal processing (TSP), pp 1–5
8. Qi J, Wang Z, Lin X, Li C (2018) Learning complex spatio-temporal configurations of body joints for online activity recognition. IEEE Trans Human-Mach Syst 1–11. https://doi.org/10.1109/THMS.2018.2850301
9. Phyo CN, Zin TT, Tin P (2019) Deep learning for recognizing human activities using motions of skeletal joints. IEEE Trans Consum Electron 65(2):243–252. https://doi.org/10.1109/tce.2019.2908986
10. Jalal A, Mahmood M, Hasan AS (2019) Multi-features descriptors for human activity tracking and recognition in indoor-outdoor environments. https://doi.org/10.1109/IBCAST.2019.8667145
11. Ehatisham-ul-Haq M, Javed A, Azam MR, Malik HM, Irtaza A, Lee IH, Mahmood MT (2019) Robust human activity recognition using multimodal feature-level fusion. IEEE Access 7:60736–60751
12. Zhou B, Yan J, Li Q (2019) Smartphone-based activity recognition for indoor localization using a convolutional neural network. Sensors 19(3):621. https://doi.org/10.3390/s19030621
13. Zalluhoğlu C, Ikizler N (2019) Region based multi-stream convolutional neural networks for collective activity recognition. J Vis Commun Image Representation 60. https://doi.org/10.1016/j.jvcir.2019.02.016
14. Long J, Sun W, Yang Z, Raymond OI (2019) Asymmetric residual neural network for accurate human activity recognition. Information 10(6):203. https://doi.org/10.3390/info10060203
15. Gnouma M, Ejbali R, Zaied M (2020) Video anomaly detection and localization in crowded scenes. In: Martínez Álvarez F, Troncoso Lora A, Sáez Muñoz J, Quintián H, Corchado E (eds) International joint conference: 12th international conference on computational intelligence in security for information systems (CISIS 2019) and 10th International conference on European transnational education (ICEUTE 2019). CISIS 2019, ICEUTE 2019. Advances in intelligent systems and computing, vol 951. Springer, Cham
16. Dhiman C, Vishwakarma DK (2019) A robust framework for abnormal human action recognition using $\mathcal{R}$-transform and Zernike moments in depth videos. IEEE Sens J 19(13):5195–5203. https://doi.org/10.1109/jsen.2019.2903645
17. Najar F, Bourouis S, Bouguila N, Belghith S (2018) Unsupervised learning of finite full covariance multivariate generalized Gaussian mixture models for human activity recognition. Multimedia Tools Appl 1–23

18. Moencks M, De Silva D, Roche J, Kondoz A (2019) Adaptive feature processing for robust human activity recognition on a novel multi-modal dataset
19. Gowda SN (2017) Human activity recognition using combinatorial deep belief networks. In: 2017 IEEE Conference on computer vision and pattern recognition workshops (CVPRW), Honolulu, HI, pp 1589–1594. https://doi.org/10.1109/CVPRW.2017.203
20. Cherian A, Fernando B, Harandi M, Gould S (2017) Generalized rank pooling for activity recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3222–3231

Image Compression Using Vector Quantization

Rohit Agrawal and Narendra Mohan

Abstract Pictures are changed over into computerized information by changing over every one of the pixels in the picture. The more hues that are utilized, the more noteworthy the measure of information that is required to store the picture. Picture pressure is the way toward diminishing these number of bits required to speak to a picture. The mainstream picture pressure calculation is the vector quantization, which maps the pixel force vectors into twofold vectors ordering a predetermined number of potential multiplications. The picture is divided into numerous squares, and each square is considered as a vector. After quickly checking on the central thoughts of vector quantization, we present a technique for the codebook structure for vector quantization calculations that perform picture handling.

Keywords Image compression · Quantization (VO) · Encoder · Decoder · Codewords · Codebook

1 Introduction

Picture compression is the planning of an informational index into a piece stream to diminish the quantity of bits required to speak to the informational index by disposing of repetitive data about neighboring pixels. With information compression, one can store more data in a given extra room and send data quicker over correspondence channels. The two sorts of information compression are lossless and lossy [1–5].

Lossless compression has the favorable position that the first data can be recuperated impeccably from the packed information. Short lists are appointed to high likelihood images, and long records are doled out to low likelihood images. The compression

R. Agrawal (✉) · N. Mohan
Department of Computer Engineering and Application, GLA University, Mathura, India
e-mail: rohit.agrwal@gla.ac.in

N. Mohan
e-mail: narendra.mohan@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_34

proportion is restricted by the source entropy rate, a proportion of the arbitrariness innate in the source [6–10].

Lossy compression brings blunder into the information, so the first information cannot be entirely recouped. Vector quantization (vector quantizes) depicted in the following segment is a calculation for the lossy compression. A picture compression framework may fill in as a front end for a computerized picture processor. Advanced picture handling is the preparing of a two-dimensional arrangement of information [11–14]. It incorporates portrayal and displaying, upgrade, rebuilding, investigation, and remaking. Pictures are regularly handled in various manners, and there are open inquiries concerning how the picture preparing activities connect with one another. Which activity ought to be performed first? Which activity makes another preparing task more straightforward or more mind boggling? Two traits of vector quantizer recommend that such shrewd compression is conceivable [15–19]. To start with, vector quantizer structure procedures of bunching and grouping trees have a long history of utilizations to picture preparing, including upgrade and order. Second, since a vector quantizer framework utilizes an assortment of conceivable picture generation hinders, a picture preparing routine can be applied to this arrangement of squares early as opposed to the compacted picture itself [20–22]. By joining other sign preparing objectives into the structure of vector quantized, the compression framework can be better redone for a specific application [23].

Images often get distorted by noise during identifying image and/or transmitting in any image application. De-noise technique thus becomes significant problem in prior processing of images. Many image de-noise techniques have recently been suggested to track impulse noise and seek to save information of images [1–4]. Many of those use the regular median filter, or its modifications or generalizations because of their pleasing ability to eliminate impulse noise. Such methods, however, typically process the entire image to change both types of pixels with and without noise [5–9]. The switching technique was implemented in recent years to prevent the impact on noise-free pixels. The switching median filter has of two steps: (i) impulse detection and (ii) noise filtering [10–15]. The key benefit of flipping median filter is that by using an impulse detector to find and filter correct noisy pixels, it can prevent damage to noise-free pixels. Prior impulse detectors, however, also need a lot of computations [16–19]. Many of them require duplications in their applications, and wide computations and broad memory space are needed [19–22].

2 Vector Quantization

Vector quantization (vector quantizer) is a compression technique for loss information, depending on the square coding guideline. This is a measure of the fixed to fixed lengths. Due to the need for multi-dimensional joining, the design of a vector quantizer (vector quantizer) is regarded as a difficult problem in the previous days.

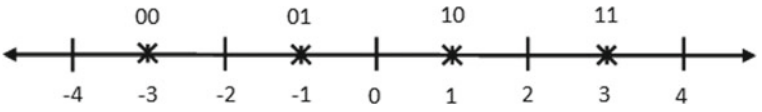


Fig. 1 1D vector quantizer

Linde, Buzo, and Gray (LBG) proposed a method for measurement of vector quantizers based on a planning structure since 1980 [24]. The use of a grouping of plans sidesteps multi-dimensional reconciliation criterion.

2.1 Preliminaries

A vector quantized is approximated and is comparable to that "rounding-off" (means to closest integer), e.g., of a 1D vector quantizer as shown in Fig. 1.

Any number not exactly -2 here is approximated to -3. Increasing in digit from -2 to 0 is approximate to -1. Each digit is approximated by $+1$ somewhere in the range 0 and 2. Every more prominent number than 2 is approximate to $+3$. Notice that exceptional 2-bits are spoken to surmised qualities. It is a quantizer of a one-dimensional two-piece matrix. It has a 2-bit pace/measurement.

A case of a quantizer of two-dimensional vectors is shown as follows. Here, a star linked to that district approximates each pair of numbers dropping at a particular locale. Notice that there are 16 localities and 16 planets—each of which can be spoken to by 4-bits exceptionally. Hence, this is a quantizer of two-dimensional 4-bit vectors. The rate is exactly 2-bits/measure.

In both examples above, that stars are known as code vectors and by borders, encoding regions are known as regions identified. The set of all code vectors is known as the codebook, and set is known as space partition of all encoding regions as shown in Fig. 2.

2.2 Design Problem

Problem with plan vector quantizer can be expressed as follows. Because of vector source with its known factual characteristics, given twisting scale, and given number of code vectors, discover codebook (arrangement of each red star) and a parcel (arrangement of blue lines) that causes the slightest normal distortion. We accept that preparation grouping comprises M source vectors:

$$T = \{x_1, x_2, \ldots, x_M\}.$$

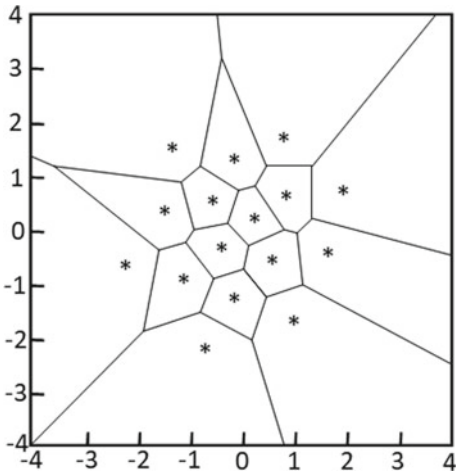


Fig. 2 Two-dimensional, 4-bit vector quantizer

This arrangement for the planning can be obtained from some massive database. For example, in case source is a discourse signal, planning succession can be obtained at that point by recording a couple of long phone discussions. M is thought to be adequately huge with goal that preparation arrangement would catch all observable properties of source. We accept that k-dimensional are source vectors, e.g.,

$$\text{“}X_m = \left(x_{m,1}, x_{m,2}, \ldots, x_{m,k}\right), \quad m = 1, 2, \ldots, M.$$

Let N be the number of code vectors, and let $C = \{C_1, C_2, \ldots, C_N\}$ represent the codebook. Each code vector is k-dimensional, e.g.,

$$c_n = \left(c_{n,1}, c_{n,2}, \ldots, c_{n,k}\right), \quad n = 1, 2, \ldots, N.$$

Let S_n be encoding region associated with code vector C_n, and let $P = \{S_1, S_2, \ldots, S_N\}$ denote the partition of the space. If the source vector x_m is in the encoding region S_n, then its approximation (denoted by $Q(x_m)$) is C_n:

$$Q(x_m) = C_n, \quad \text{if } X_m \in S_n.\text{”}$$

If average distortion is given, let a squared error distortion measure be

$$D_{\text{eve}} = \frac{1}{Mk} \sum_{m=1}^{M} ||X_m - Q(X_m)||^2$$

where

$$||e||^2 = e_1^2 + e_2^2 + e_k^2$$

Given T and N, find C and P such that D_{ove} is minimized. Design problem can be succinctly stated as follows.

2.3 Optimality Criteria

To the above minimization problem, if C and P are a solution, then it must satisfy following two benchmarks.

Nearest Neighbor Condition

$$\text{“}S_n = \left\{x : ||x - C_n|| \leq ||x - C_{n'}||^2 \forall_{n'} = 1, 2, \ldots, N\right\}\text{”}$$

This condition says encoding district S_n will consist of all vectors similar to C_n than any of various code vectors. Any tie-breaking strategy would do to those vectors lying on top.

Centroid Condition

$$\text{“}\boldsymbol{c_p} = \frac{\sum_{x_m \in S_n} X_m}{\sum_{x_m \in S_n} 1} \quad n = 1, 2, \ldots, N\text{”}$$

This condition says code vector c_n should be standard for each of those preparation vectors that are in district S_n encoding. In use, one can guarantee that one preparing vector has a position with each encoding district at any rate (thus, denominator in above condition is rarely 0).

Quantization

A vector quantizer maps k-dimensional vectors in the R_k space into a finite set of vectors $Y = \{y_i: I = 1, 2, \ldots, N\}$—vector y_i is called a code vector or codeword, and a codebook is called the set of all the codewords. Associated with every codeword, y_i is a competitor region called Voronoi and is denoted as

$$\text{“}V_i = \left\{x \in R^k : ||x - y|| \leq ||x - y_j||, \text{ for all } j \neq i\right\}\text{”}$$

The set of Voronoi regions partition the entire space R' such that

$$\text{“} \cup_{f=1}^{N} V_i = R^k\text{”}$$
$$\text{“} \cap_{i=1}^{N} V_i = R^k\text{”}$$
$$\text{“for all } i \neq j\text{”}$$

Or in two-dimensional cases, we take vectors without loss of all-inclusive argument. Figure 3 mentions certain space vectors. A delegate codeword is associated

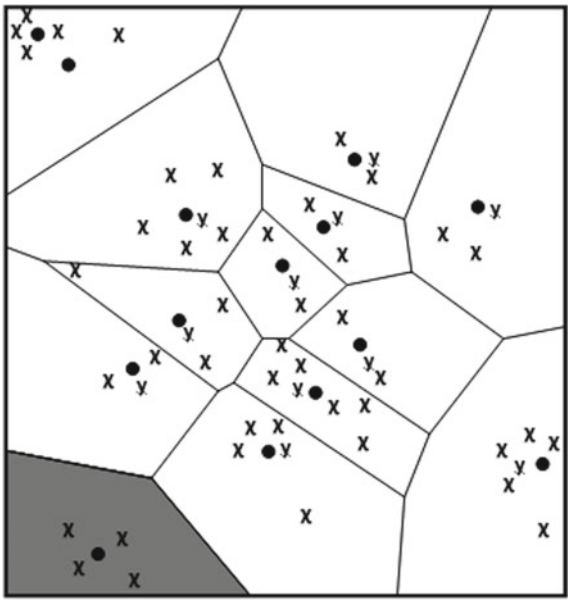


Fig. 3 Codewords in two-dimensional space (x is marked as input vectors, codewords are marked with red circles, and Voronoi regions are separated with boundary lines)

with each bunch of vectors. Each codeword resides in its own locality Voronoi. For outline, these assumptions are isolated with zero lines in Fig. 3. In the light of a data matrix, the codeword chosen to refer to it is in one in equivalent Voronoi district.

To be closest in Euclidean distance, representative codeword is determined from input vector. Euclidean distance is denoted by

$$d(x, y_i) = \sqrt{\sum_{j=1}^{k} \left(x_j - y_{ij}\right)^2}$$

where jth component of input vector is x_i and jth is component of codeword y_i is Y_{ij}.

3 Vector Quantizer in Compression

The composed two undertakings are vector quantized. Encoder is the first element in vector quantizer, and decoder is the second element in vector quantizer. Encoder takes information from vector and yields record of codeword that opportunity most decreased mutilation. By evaluating Euclidean partition for this circumstance, least turning is found between information vector and every codeword in codebook. Once closest codeword is discovered, document of that codeword is sent through a channel (could be communications channel, computer accumulating, and so on). When encoder gets rundown of codeword, it replaces document with related codeword.

3.1 Search Engine Working in Encoder

In spite of certainty that as scalar quantization, vector quantizer opportunity more compression for a comparable winding rate and PCM, yet is not as commonly executed. As a result of two things, time is the first element as it gets to make codebook, and speed of chase is the second element. Various algorithms have been proposed to enhance the speed of the search. To choose codeword, some of them reduce math utilized that offers base mutilation; various computations preprocess codewords and experience essential structure.

The least difficulty request approach, which is similarly slowest, has the whole interest. In whole chase, information vector is proportioned with each codeword in codebook. On remote possibility that M is input vectors, N is codewords, and every vector is k estimations, by then amount of increments occurs kMN, amount of increases and reasonings become "$MN((k-1)+k) = MN(2k-1)$", and amount of proportions becomes "$MN(k-1)$". This prepares whole request a costly technique.

Measure of Performance

There is nothing but bad approach to gauge the exhibition of vector quantizer. This is because the distortion that VQ incurs will be evaluated by us humans and that is a subjective measure. Try not to surrender! We can generally fall back on old fashion, and the ratio between mean squared error (MSE) and peak signal-to-noise ratio (PSNR) is obtained. MSE is characterized as follows:

$$\text{MSE} = \frac{1}{M}\sum_{i=1}^{M}\left(\widehat{x_i} - x_i\right)^2$$

where M is quantity of components in sign or picture. For instance, in event that we needed to discover MSE between reconstructed and first picture, at that point we would receive distinction between two pictures resolution by resolution, square outcomes, and normal outcomes. PSNR is characterized as follows:

$$\text{PSNR} = 10\log_{10}\left(\frac{(2^n-1)^2}{\text{MSE}}\right)$$

where n is quantity of bits per image. For instance, on off chance that we need to discover PSNR between two 256 dim-level pictures, at that point we set n to 8-bits. The presentations of vector quantizer are regularly given regarding the sign to bending proportion (SDR):

$$\text{SDR} = 10\log_{10}\frac{\sigma^2}{D_{\text{ove}}}(\text{in dB}),$$

where

σ^2 variance of source and

D_{ove} average squared error distortion. The higher sign to bending ratio has better performance.

4 Conclusion

Advanced picture compression is a significant strategy in computerized picture handling. To improve its presentation, we might want to accelerate the structure procedure and accomplish the most noteworthy compression proportion conceivable. Toward this path, our objective has been to depict the major thoughts of vector quantization and to present a fundamental codebook structure strategy dependent on precise order and association of square examples. The recently created codebook structure technique helps in proficient picture compression.

References

1. Gersho A, Gray RM (1991) Vector quantization and signal compression. Kluwer, Boston, MA
2. Linde Y, Buzo A, Gray RM (1980) An algorithm for vector quantizer design. IEEE Trans Commun COMM 28:84–95
3. Neal JBO (1976) Differential pulse with entropy coding. IEEE Trans Inf Theory IT 22:169–174
4. Kumar A, Sharma K, Dixit AR, A review of the mechanical and thermal properties of graphene and its hybrid polymer nanocomposites for structural applications. J Mater Sci 54(8):5992–6026
5. Sharma K, Shukla M (2014) Three-phase carbon fiber amine functionalized carbon nanotubes epoxy composite: processing, characterisation, and multiscale modeling. J Nanomater
6. Kohonen T (1988) Self-organization and associate memory, 2nd edn. Springer, Berlin, Germany
7. Sharma K, Kaushalyayan KS, Shukla M, Pull-out simulations of interfacial properties of amine functionalized multi-walled carbon nanotube epoxy composites. Comput Mater Sci 99:232–241
8. Yadav A, Kumar A, Singh PK, Sharma K, Glass transition temperature of functionalized graphene epoxy composites using molecular dynamics simulation. Integr Ferroelectr 186(1):106–114
9. Singh PK, Sharma K, Kumar A, Shukla M, Effects of functionalization on the mechanical properties of multiwalled carbon nanotubes: A molecular dynamics approach. J Compos Mater 51(5):671–680
10. Singh PK, Sharma K, Mechanical and viscoelastic properties of in-situ amine functionalized multiple layer graphene/epoxy nanocomposites. Curr Nanosci 14(3):252–262
11. Vaisey I, Gersho A, Image variable block size segmentation. IEEE Trans Signal Process ASSP 40:2040
12. Singh PK, Sharma K (2018) Molecular dynamics simulation of glass transition behaviour of polymer based nanocomposites. J Sci Ind Res 77(10):592–595
13. Kumar A, Sharma K, Dixit AR, Carbon nanotube-and graphene-reinforced multiphase polymeric composites: review on their properties and applications. J Mater Sci 1–43
14. Shukla MK, Sharma K, Effect of carbon nanofillers on the mechanical and interfacial properties of epoxy based nanocomposites: a review. Polym Sci Ser A 61(4):439–460
15. Kumar A, Sharma K, Dixit AR, A review on the mechanical and thermal properties of graphene and graphene-based polymer nanocomposites: understanding of modelling and MD simulation. Mol Simul 46(2):136–154

16. Mausam K, Sharma K, Bharadwaj G, Singh RP, Multi-objective optimization design of die-sinking electric discharge machine (EDM) machining parameter for CNT-reinforced carbon fibre nanocomposite using grey relational analysis. J Braz Soc Mech Sci Eng 41
17. Yang SB, Tseng LY (2001) Smooth slide classified vector quantizer with variable block size. IEEE Trans Image Process 685
18. Kumar K, Sharma K, Verma S, Upadhyay N, Experimental investigation of graphene-paraffin wax nanocomposites for thermal energy storage. Mater Today Proc 18:5158–5163
19. Goyal M (2018) Shape, size and phonon scattering effect on the thermal conductivity of nanostructures. Pramana 91(6):87
20. Ziv J, Lempel A (1977) A universal algorithm for sequential data compression. IIEEE Trans Inf Theory IT-23(3):337–343
21. Goyal M, Gupta B (2019) Study of shape, size and temperature-dependent elastic properties of nanomaterials. Mod Phys Lett B 33(26):1950310
22. Goyal M, Gupta B (2019) Analysis of shape, size and structure dependent thermodynamic properties of nanowires. High Temp High Pressures 48
23. Goyal M, Singh M (2020) Size and shape dependence of optical properties of nanostructures. Appl Phys A 126(3):1–8
24. Agarwal R, Jalal AS, Arya KV (2020) A review on presentation attack detection system for fake fingerprint. Mod Phys Lett B 34(05):2030001

Adaptive Bat Optimization Algorithm for Efficient Load Balancing in Cloud Computing Environment

Rakesh Kumar, Diwakar Bhardwaj, and Ranjana Joshi

Abstract Cloud computing is the new development technology where it permit users as much as needed with its greater performance. Where the cloud computing is a heterogeneous method where it contains huge number of various system data. With this system tasking few data else calculating with greater systems, where it accepts the computing that the forwarding and accessing time is very difficult with system. In cloud environment, the load managing is the important function to get performed. Where the load balancing permits to share various workloads through the various nodes and it verifies by checking that there is no individual node gets over loaded. In previous applications, where the response time is high and then the communication cost is maximum and thereby to get out of previous declined problem, with the discussed methods, and for capable load balancing, the adaptive bat optimization algorithm (ABOA) is discussed. The discussed various servers huge balancing and depending upon the independent and dependent function. The task gets at the time of the run time through the server with different approximate gap with different load specifications. Variation inequality (VI) issues are gets eliminated then there is a Nash equilibrium method fixes through the calculated game. After discussing, an iterative proximal (IP) algorithm through ABOA to get calculated a Nash equilibrium result through the various servers huge managing one. By the outcome, the discussed ABOA gives good task than that of the previous system.

Keywords Load balancing · Adaptive bat optimization (ABOA) · Cloud computing

R. Kumar (✉) · D. Bhardwaj · R. Joshi
Department of CEA, GLA University, Mathura, India
e-mail: rakesh.kumar@gla.ac.in

D. Bhardwaj
e-mail: diwakar.bhardwaj@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_35

1 Introduction

With the huge area forwarding calculations, the cloud computing enters as the new paradigm. It gets forwarded computing then the data apart through the desktop gets transfer PCs, with the maximum data areas. Cloud computing gives the fine IT sourcing like services and applications, and then they functions on the fine surroundings, with the Internet areas, with pay-per-usage depending to manage quantity suddenly and comfortably. Where it functions to accompany the modifications through manages then it promote with various industrial by ignoring the amount of hardware and software. Hence through cloud computing method is used for operates a apt, with the demand network manages with the partition pool through the calculating resources (e.g., servers, storage, services, network, and applications). These cloud computing helps in de provisioned and provisioned immediately through less handing methods else the service managing method. Thereby, it encourages capability. Because of the maximum growth in the cloud computing, they have been whole occupied through the industry then they have quick elaboration in the data centers.

Cloud computing is the one for establishing all around, with mode demanding and quick accessible are get transferred through the point of pool improvement in the cloud solutions (e.g., storage, network, servers, services, and then the application) through less management can be easily provisioned and transmitted else through the service interaction provider. Promoting the similar resources with the cloud loads based on the quality of service (QoS) needed by the workloads. With the higher usages with the cloud, the cloud services have become in higher amount depending upon the quality mainly higher problem due to various huge tasks where it gets proposed individually and gets comparing to trust and then the possibility. Where the cloud computing is the huge IT depending surrounding based on the usages and has made the location where it gets served scalable based with the demand, metered services to the users of the cloud and then the elastic. Through this one gets huge spotted in dual individual and leveled.

Because of the capability, where it assures the users quality of services (QoS) and finding the utilization of availability to get higher functions, how to get spaced radio resources higher and as usual the hot topic for the upcoming wireless communication. For getting the problem to be resolved, the resource allocation is estimated and the issue gets solved in online with sudden instance information. Where many allotment issues get convex, the real time the fine resultant with huge difficult one. Then the optimal systems are allotted through facing higher QoS needs through the radio resources, where improves the traverse through the well planned resource allotments. It permits three important methods: It offers three primary services: infrastructure-as-a-service (IaaS), platform-as-a-service (PaaS), and then the software-as-a-service (SaaS). With the SaaS, greater cloud gives the sponsor software products where the users beyond the Internet. By this application information could be processed at whenever and at any locations within the devices. Within the PaaS, cloud user helps in giving the surroundings to make and relay custom systems; thus, the users not to worry regarding the infrastructural needs also have to handle through the systems and

getting formed and created. With the IaaS, execute the functions and get saved where they sponsor at a fee. Resources like the memory, network bandwidth, CPU, storage are created and capable to many exchanging users. Where these sharing of polled methods are many users called multitenancy. Where the cloud deployment model consists of four and they are public, hybrid clouds, private, and then the community. Where the private clouds are taken through the organization, then the only one staff gets allotted to get handled the private cloud. Then, it as usual handled through the in-house specialist else out-sourced to other parties.

One of the problems in the cloud execution environment is the load balancing reason. The important aim of the load balancing is to get elaborating the higher resource utilization, extend the utilization of resources, deprecation reaction time, and accessible scalability, eliminating the not needed burden of any such assets and ignoring bottlenecks. In this system, the number of loaded works in the system counting through elaborating methods performance and the usage of the resources. Within the cloud computing load, there will be the energy memory availability, CPU limitations, restricted memory capacity. Hence, all over load balancing needs not be at certain conditions where the single server else the information center is below use and another getting loaded mode. The improvement algorithm has involved through getting finite load balancing with the cloud environment.

In cloud environment, there are only simple works on the literature. Tsai et al. noticed to finite task allocating and resource allocation are done through the enhanced variation evolution algorithm (IDEA) depending on the discussed amount and time models through cloud computing environment [1]. Vakilinia et al. proposed through managing the optimization allocating and calculated methods; in data centers (DC), the power consumption gets minimized. Shao et al. handle the load balancing method depending on data correlation with the cloud computing [2]. Velde et al. educate the theory load balancing algorithm utilization of fuzzy methods with in cloud computing, where the load balancing is a average and difficult issue within cloud computing [3]. Gupta et al. pick a meta-heuristic method of ant colony optimization (ACO) algorithm to getting resolved the function scheduling issues in cloud environment depends importantly with the two aspects, i.e., less the makespan/executing time then the good load balancing [4, 5]. In [6–9], authors suggest other applications of load balancing also.

2 Proposed Methodology

With this work, adaptive bat optimization algorithm (ABOA) is discussed to enhance the load balancing through server accessible model efficiently through the cloud environment.

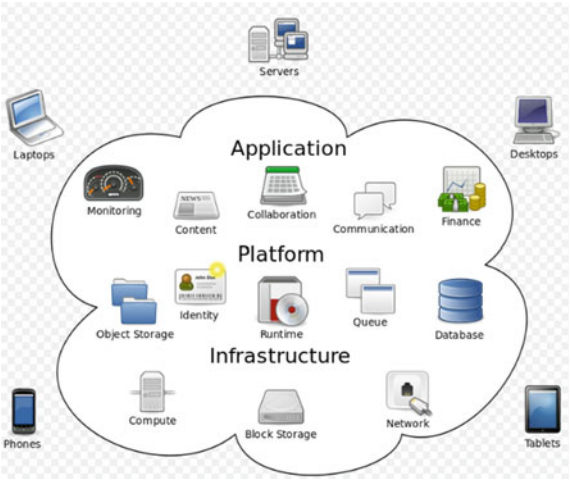


Fig. 1 Cloud computing model

2.1 System Model

Getting start through the current system method within the context of the cloud promotes through the m heterogeneous servers, where it gets joints through the communication network. It specifies server set as $M = \{1, \ldots, m\}$, the server processing capacity $i\,(i \in M)$ is indicated through the service rate μ_i, where it is also affected through the workload within the server. At the server, the external request arrival rate is $I\,(i \in M)$ has taken to start follow a Poisson system. By taking into with every servers, the information systems are well defined as $M/M/1$ queuing method [10] (Fig. 1).

2.2 Request Migration Model

By taking into consideration, the request migration model at the getting migration profile at the server $i(i \in M)$ is calculated in the below chances vector.

$$p_i = (p_{i_1}, \ldots, p_{i_m})^{\mathrm{T}} \tag{1}$$

Then, $p_{ij}(j \in M)$ is the random one where it demand at the server i is moves to the server j then it leads to some limitation $\sum_{j=1}^{m} p_{ij} = 1$. Where the limitation forwarding method set of the server $i\,(i \in M)$ is defined as

$$P_i = \left\{ p_i \left| \sum_{j=1}^{m} p_{ij} \text{ and } p_{ij} \geq 0, \quad \forall j \in M \right. \right\} \tag{2}$$

There are approximate reached external demands with the every server.

Specified ϕ_i in the external demand reached amount with the server $i(i \in M)$, then the λ_{ij} is the demand flow amount through server i to server j (i.e., calculated amount of demands is waiting amount of demand transfer through the server i to server j per unit of time). After,

$$\lambda_{ij} = p_{ij}\phi_i \tag{3}$$

Then, the demand migration profile is be defined as

$$= (\lambda_{i_1} \ldots \lambda_{i_m})^{\mathrm{T}} \tag{4}$$

Similarly, the demand migration method fix is defined as

$$Q_i = \left\{ \lambda_i \left| \sum_{j=1}^{m} \lambda_{ij} = \phi_i \text{ and } \lambda_{ij} \geq 0, \quad \forall j \in M \right. \right\} \tag{5}$$

Then, ϕ_i is the external demand reaching amount at server I then it mingles with every single method and fix with each servers are defined by

$$Q = Q_1 \times \cdots \times Q_M. \tag{6}$$

Specifying μ_j as the execution rate with the server $j(j \in M)$. After, the approximate demand at every server not be greater its executing levels, i.e., $\beta_j = \sum_{i=1}^{m} \lambda_{ij} < \mu_j$, then β_j is the loaded with work of server $j(j \in M)$. After, by providing the different demand forwarding methods $\lambda_{-i} = \left(\lambda_j\right)_{j=1, j\neq i}^{m}$, the demand forwarding method λ_i of server $i(i \in M)$ then content $\sum_{i=1}^{m} \lambda_{ij} < \mu_j$.

2.3 *Independent and Dependent Task Using Server Availability*

With this analysis, dependent and independent functions are made through the provided server accessible method. The structure of the discussed methods is well organized through the load balancing through the cloud and load independent server and load dependent server possibility. Through the supporting-based relation through the systems gets aligned through the users in the system of the cloud computing tasking and comes to the boosting one. Dependent functions (for example, single work may depend with one another to attain the targeting commencement) from users taking the dynamic needs, where it gets necessary to fine the cloud usages through the path where the power gets occurred and attains the less one. Where the

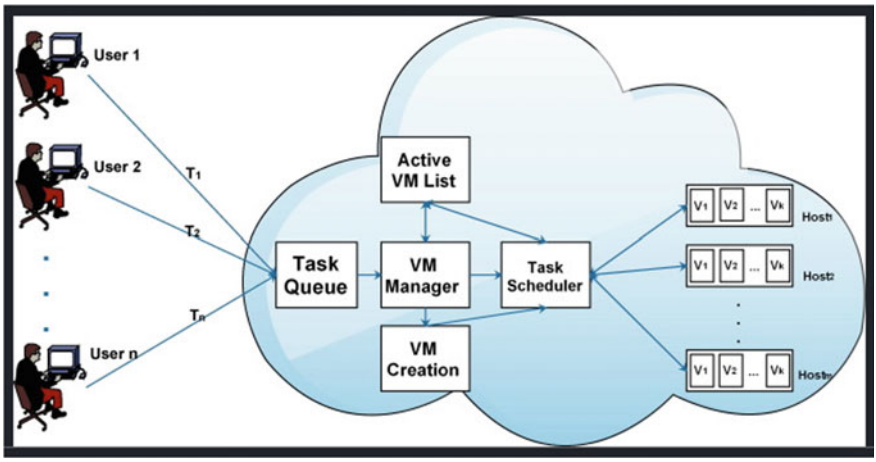


Fig. 2 Task scheduler

dependent task performs the task where the users make the submitted of the dependent relationship with each other. When the tow functions get fixed, then there is dependent bonding in between them, where the subtask gets pointed toward each other with the parent task. The cost amount gets averaged and by getting gathered with every systems of task, then the approximate estimation amount indicates the average amount of the cost with the task server in overall server resources. With the same executing node, the tow dependencies are transferred, then there is no transferring costing in between on another (Fig. 2).

With the dependent function, reverses with overall system then estimates the preference with every system, where the system gets out of in the lower to higher order. Where it estimates the processing amount of the task with every executing method and then appoints the interval time of the task. Searching all parent task, then it gets solved. Replicating the parent task then it gets estimated system and then appoints the waiting time to get proceeded. At the end, by estimating the entire time of the task and then upgrade the sudden time. Server accessibility and it gets related with randomly where the server noticed that the server has noticed with the executing state and the providing point with the time. Also it gets explained and the calculation of the time where the executing is providing the same time gap. The bonding in between the approximate load intensity and get defected amount not permitted. Where the workload intensity gets maximized, then higher defects are noticed. Based on the present load index, the current load executions with overall systems are transferred through the cloud server machines.

2.4 Load Balancing Using ABOA Algorithm (LB-ABOA)

Using ABOA algorithm, the load balancing is done. With this area, where it gets calculated the inquiry migration issues in the many servers as well as non-cooperative game. Through different not equality (VI) theory, it finds the presence of Nash equilibrium results to set through calculated game. It gets discussed through variation proximal algorithm to calculate a Nash equilibrium results.

Where the load balancing has importantly played with its task in cloud computing where it load balancing and targeting to elaborate resource capturing, by getting more output, limit response time, then ignores overload of various individual resource. Good load balancing helps cloud execution higher capability and enhances user needs. Thereby, it is the system of promising the same distribution of work load with the pool of method of nodes else the operator; hence, the functioning systems are evenly transmitted of work load with the pool of method node else the system; hence, the functioning system gets accomplished not by any troubling. The aim of load balancing has to attain the stable of system, enhance the improvement, create the method, where it get defected and providing differentiation with the function like safety up gradation, exhibit the customers resource, and time through upcoming function system. Cloud load is one type of load balancing system where it is processed with the cloud execution where it gets fulfilled separately based on the cluster dependencies (Fig. 3).

Bats are attractive animals with their approach ability of echolocation has been fascinated to pay of various researchers through various areas. Through the old age data of cloud user's methods, the bats can execute how much distance they are from. They can able to get variations like selected benefits (food) and the non-selected benefits (obstacle) through its fabulous orientation systems. By this method, they are get created to respond as numerous bats following good benefits of old data through the ability of echolocation. Below one are the overall diagrammatic representation and shown in Fig. 4.

1. For capturing location and how much distance area bats utilizes the echolocation, as well as they also know the variations through the olden data with the cloud allocation in different extraordinary methods;
2. The bat b_i flew approximately through the velocity v_i with its location x_i at the point of frequency $f_{\min}$, different wavelength λ, then the loudness A_0 helps to finding the good benefits of olden data. Through that they independently manages the wavelength (else the frequency) with their leasing pulses then alters the amount of pulse exhibited $r \in [0, 1]$, based on the random with the classification;
3. Even though the frequency might differentiate with various methods, Yang [11] considering the high frequency differs through the huge (positive) A_0 by less number of same amount with $A_{\min}$.

Algorithm 1 describes bat algorithm (taken through [11]):

Primarily, starting point x_i, velocity v_i then the frequency f_i operates through every bat b_i. With every time count t is get T with the higher amount of variations,

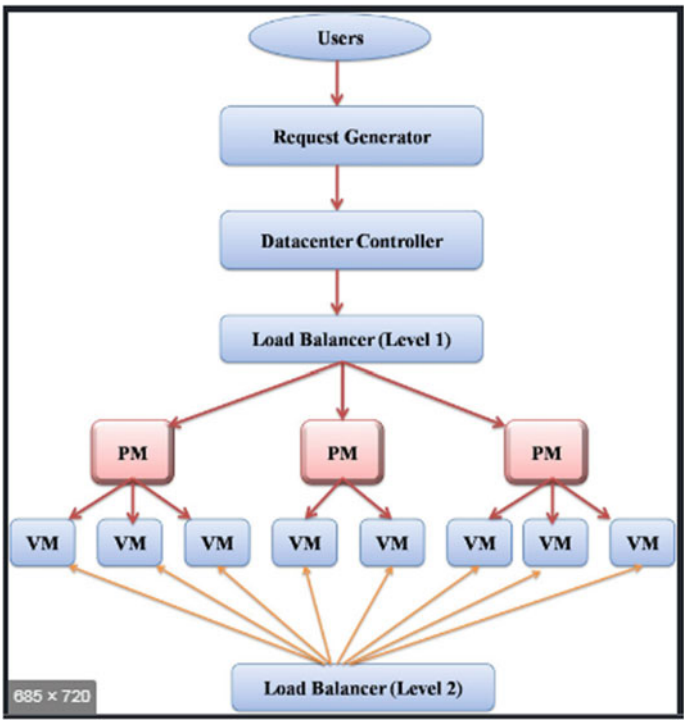


Fig. 3 Load balancing

the virtual bats motions are overloading through its position and the velocity through the Eqs. 7, 8, and 9, as below:

$$f_i = f_{\min} + (f_{\min} - f_{\max})\beta, \tag{7}$$

$$v_i^j(t) = v_i^j(t-1) + \left[\hat{x}^j - x_i^j(t-1)\right] f_i, \tag{8}$$

$$x_i^j(t) = x_i^j(t-1) + v_i^j(t), \tag{9}$$

By β specifies the variation created amount through the meantime [0, 1]. Calling $x_i^j(t)$ specifies through indicating the amount of considering changing the values j the bat i with its time count t. The output f_i (Eq. 2) helps in handling the range and pace with the rotations of in the bats. The changing $\hat{x}^j$ presents the present global good chosen benefits of location (solution) through taking decided variable j gets attained through distinguish with every choosing benefits of the resultant by given m bats. The good resultant gets formed with the new local resultant which gets formed. Where these resultants get formed through the variable walk through the below method: Exploitation and the exploration are gets balanced, Instead of making the

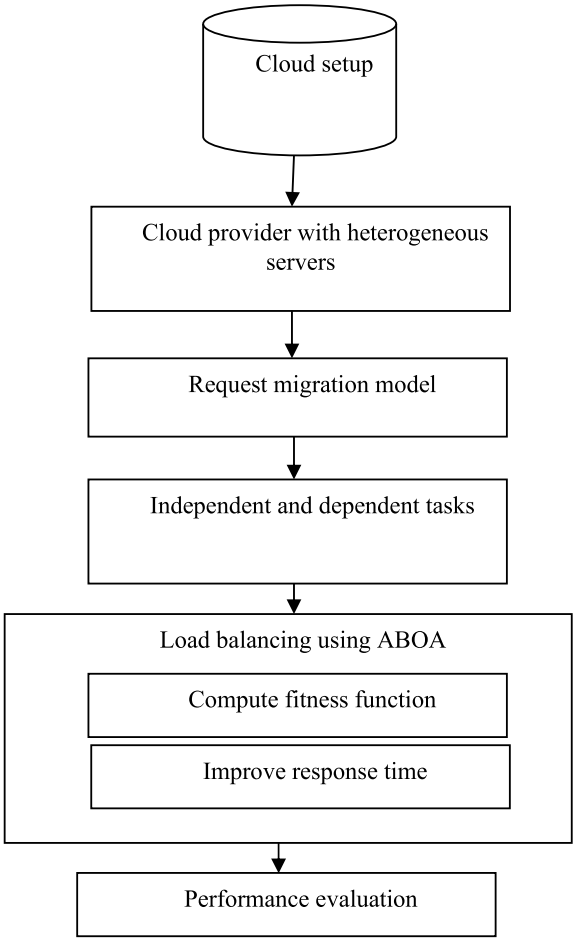


Fig. 4 Overall block diagram of the proposed system

variable walk, the adaptive repetition step size is utilized to create new resultant. By enhancing the varieties of the chances of choosing the benefit of the resultant, it helps to employ variable walks. Importantly, single choosing benefits resultant id picked through the present good benefits resultant, then the approximate walk gets implemented by creating the new benefits picking resultant picking solution through every bat where it agrees the restrictions on the Line 5 of Algorithm 1:

$$x_{\text{new}} = x_{\text{old}} + \varepsilon \overline{A}(t) \tag{10}$$

where the $\overline{A}(t)$ specifies the approximate loudness among every bats with the time t, then the $\varepsilon \in [-1, 1]$ targeting to the strength and direction with the oscillation walk. For every repetitions of the algorithm, loudness A_i then the transmitting pulse rate r_i gets uploaded, as below one:

$$A_i(t+1) = \alpha A_i(t) \tag{11}$$

Then,

$$r_i(t+1) = r_i(0)\big[1 - \exp(-\gamma t)\big], \tag{12}$$

The ad hoc stability is the α and γ. In the algorithm first count, the transmission rate $r_i(0)$ then the loudness $A_i(0)$ gets usual approximate picked one. In general, $A_i(0) \in [1, 2]$ then the $r_i(0) \in [0, 1]$. In the new bat methodology, the Gaussian distribution is used. Gaussian distribution specifically creates the fewer amounts but randomly creates higher amount too. Two methods are described in the distribution: mean μ then the steady deviation σ. The degree with the minimum amount through the maximum ones gets handled through easy modification with fixed deviation σ in the distribution. Thereby, the fresh bat algorithm handles the approximate walk count size through the difference of Gaussian/normal distribution. The changing equation is below:

$$x_{\text{new}} = x_{\text{old}} + \varepsilon \overline{A}(t) N(0, \sigma) \tag{13}$$

Algorithm 1 Load Balancing-Adaptive Bat Optimization Algorithm (LB-ABOA)

Objective function $f(x), x = (x^1, \ldots, x^n)$.
Starting the bat counting x_i in the benefits then the $v_i, i = 1, 2, \ldots, m$.
Explain pulse frequency f_i at $x_i, \forall_i = 1, 2, \ldots, m$.
Starting pulse rates r_i then the loudness $A_i, i = 1, 2, \ldots, m$.

1. Where $t < T$
2. With every bat b_i, as the attribute do
3. Create new resultant with the Eqs. (7), (8), and (9).
4. If rand $> r_i$, then
5. Choose features through the good resultant.
6. Create a local solution over the good resultant.
7. If rand $> A_i$ and $f(x_i) < f(\hat{x})$, then
8. Takes the new choosing benefit resultant.
9. Implement adaptive bats through enhancing comeback with the cost and the time through (13).
10. Raising r_i and limiting A_i.
11. Rank bats then choose present good features $\hat{x}$.

3 Experimental Result

With these work, central point gets variable than that of the experimental setup, the discussed system enhances with a not cooperative surroundings and then in the well developed through the distributed ones. The response time is taken through the tasking metrics, interlinking amount then all over time issues. The previous methods like ACO and SVM algorithms have distinguished through the discussed LB-ABOA algorithm.

Through Fig. 5, we can notice the variations of reacting time through the previous ACO, SVM, and discussed LB-ABOA method. The system gets marked through the *x* axis then the response time gets marked with the *y* axis. It displays the previous ACO and SVM system to get maximum responding time though the discussed LB-ABOA gives less reacting time. Thereby, the output shows the discussed method is good improvement with the system of enhanced load balancing through LB-ABOA method on cloud computing.

Through the graphical representation in Fig. 6 can be noticed the difference of communication amount through the previous ACO, IPA, then the discussed LB-ABOA methods. The system is marked with the *x* axis then the information amount

Fig. 5 Response time

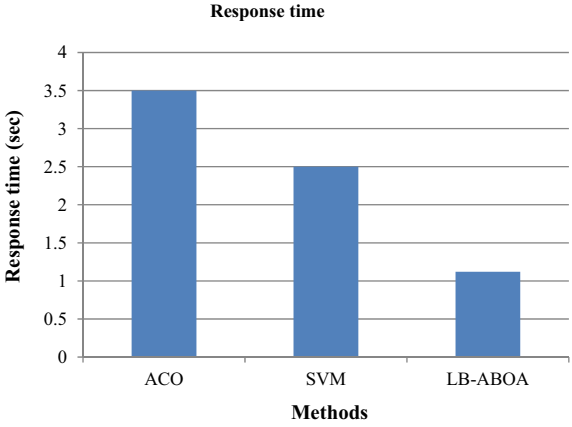
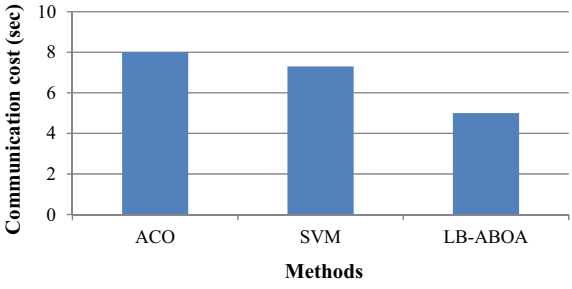


Fig. 6 Communication cost



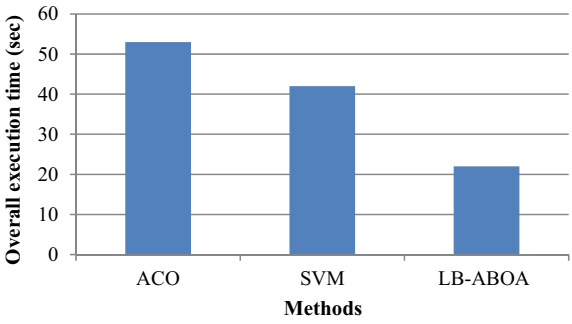


Fig. 7 Overall execution time

gets allocated with the *y* axis. Where it gets displays that the previous ACO and SVM systems give the maximum information amount though the discussed LB-ABOA gives the less communication amount. Thereby, the output gives that the discussed method that is in good improvement through the terms of enhancing load balancing through the LB-ABOA method on cloud computing.

In Fig. 7, it can notice the variation in all over processing time through the previous ACO, SVM, and discussed LB-ABOA methods. The system is allocated with the *x* axis then all over the calculating amount is allocated on the *y* axis. Where it gets displayed in the previous ACO and SVM system gives maximum processing time discussed LB-ABO gives the less all over processing time. Thereby, the output gives the discussed methods have been in good improvement through the terms of enhanced load balancing through LB-ABOA with the cloud computing.

4 Conclusion

In the business, the cloud computing has been welcomed in higher amount, even though there might be many issues such as load balancing, consolidation, virtual machine migration, and so on. This problem gets problem through load balancing, where it gets needed through transferring and gets usable dynamic local workload easy to all hubs by whole cloud to access a maximum utilized by improvement and at the portion. Through the ABOA algorithm to the multi-servers load balancing is done with the multi-heuristic methods then gets executed within this research. The discussed many servers load maintained is tasked depending on the dependent as well as the independent task. Through the LB-ABOA algorithm, the load balancing is made with higher power within the cloud computing. Thereby, the output gives the discussed methods are in good reacting amount and then the maximum improvement than that of the previous SVM and the ACO systems.

References

1. Tsai J-T, Fang J-C, Chou J-H (2013) Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm. Comput Oper Res 40(12):3045–3055
2. Shao G, Chen J (2016) A load balancing strategy based on data correlation in cloud computing. In: 2016 IEEE/ACM 9th International conference on utility and cloud computing (UCC). IEEE
3. Velde V, Rama B (2017) An advanced algorithm for load balancing in cloud computing using fuzzy technique. In: 2017 International conference on intelligent computing and control systems (ICICCS). IEEE
4. Gupta A, Garg R (2017) Load balancing based task scheduling with ACO in cloud computing. In: 2017 International conference on computer and applications (ICCA). IEEE
5. Shmilovici A (2009) Support vector machines. In: Data mining and knowledge discovery handbook. Springer, Boston, MA, pp 231–247
6. Chaturvedi A, Kumar R (2021) Multipath routing using improved grey wolf optimize (IGWO)-based ad hoc on-demand distance vector routing (AODV) algorithm on MANET. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
7. Kumar R, Chaturvedi A (2021) Improved cuckoo search with artificial bee colony for efficient load balancing in cloud computing environment. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
8. Kumar R, Bhardwaj D, Mishra MK (2020) Enhance the lifespan of underwater sensor network through energy efficient hybrid data communication scheme. In: 2020 International conference on power electronics & IoT applications in renewable energy and its control (PARC), Mathura, Uttar Pradesh, India, pp 355–359
9. Shukla DK, Dwivedi VK, Trivedi MC (2020) Encryption algorithms in cloud computing. Elsevier's J Mater Today Proc
10. Penmatsa S, Chronopoulos AT (2011) Game-theoretic static load balancing for distributed systems. J Parallel Distrib Comput 71(4):537–555
11. Yang X-S (2011) Bat algorithm for multi-objective optimization. Int J Bio-Inspired Comput 3(5):267–274

Modified Gabor Filter with Enhanced Naïve Bayes Algorithm for Facial Expression Recognition in Image Processing

Renu Prasad, Rohit Agrawal, and Himanshu Sharma

Abstract Face recognition is a subdomain of computational biometrics, which is an automated technique for verifying an identity of a human based on some parameters like the pattern in iris and face dimensions. It is a multi-faceted process as there are only a limited images available for the purpose of training, and different images are obtained through various conditions and surroundings. There is a constant challenge that researchers face which is as hoe to efficiently handle the sensitivity of the classifier under illumination, expressions, and performing under low resolution images. Conventional methods of face recognition have a lower accuracy on the classification, and hence, this paper proposes a modified Gabor filtering technique based on NB classifier (MGF + ENBA) for overcoming the issue. The MGF technique extracts only the most important features from the available data in an efficient manner. The ENBA algorithm is then used for the classification of facial expression in a more accurate way with the emotion detection such as sad, happy, angry, fear, and neutral emotional features. The experiments were performed using the Cohn-Kanade data and the JAFFE database. The results prove that the proposed technique performs better than the other methods in terms of precision, recall, and *F*-measure which prove improved accuracy.

Keywords Facial expression · Modified Gabor filter with enhanced Naïve Bayes algorithm (MGF + ENBA) · Image processing

R. Prasad (✉) · R. Agrawal · H. Sharma
Department of Computer Engineering and Application, GLA University, Mathura, India

R. Agrawal
e-mail: rohit.agrwal@gla.ac.in

H. Sharma
e-mail: himanshu.sharma@gla.ac.in

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_36

1 Introduction

Facial expression detection systems are largely being researched in the AI domain. A lot of facial expression recognition (FER) applies traditional ML methods for the feature extraction, and these traditional methods have a overhead of poor accuracy when new images are tested. Many artificial intelligence-based systems are being developed for the detection of facial expressions for improving business domains such as in customer attention marketing, health monitoring, and other intelligent systems based on emotions. Owing to a large demand and lack of supply, the facial recognition systems with emotional intelligence have attracted a large volume of research community toward the computer vision domain [1, 2].

Human emotions often facilitate more meaning rather than words. Thus, the importance of understanding the emotions is very crucial in context to the social communications. Many techniques are available for the detection of emotions such as the voice recognition, NLP, and complex techniques include EEGs. However, a much easier and stronger way for the emotion detection are through the facial expressions. The common facial expressions include anger, happiness, fear, surprise, and disgust, more complex expressions which have the combination of basic emotions are still under investigation [3, 4]. Hence, a utility which can detect the motions from facial expressions is in great demand.

The conventional method has used feature selection based on hand-crafted techniques. These include techniques such as the local binary patterns (LBPs), the non-negative factorization (NMF), and spatial learning [5, 6]. In the year 2013, various competitions on FER [7] and emotions detections are conducted which have largely collected enough training data in real-time [8, 9]. Owing to the great advancements in chip level technologies and well architected networks, various researches have commenced in the usage of deep learning methods that have achieved a lot in the accurate detection of emotions [10, 11]. Also, given the more effectiveness in the training of data, deep learning has shown a dramatic efficiency for handling the most challenging issues in the emotional detection. Authors in [12–14] suggest application areas where facial recognition is utilizes in most promising way.

In any image processing technique, feature selection plays a vital role in the classification part. If there are no sufficient features, the performance of the classifier will be poor. In most of the cases of facial expression classification, large number of features will be generated and from which only the significant features will be in need for the classification which is done through optimization techniques [15, 16]. Among the famous techniques, the Gabor wavelet feature extraction has attracted many researchers owing to its performance. Although widely adopted, the computational cost is very high for the coevolving of facial images with multi bank filters for extracting the scalar and orientation parameters. Further, this technique suffers from couple of major hindrance such as the bandwidth of the GF is always limited to a single octave and the filters are not optimized. These needs the intervention of log form in GF filters for an efficient feature extraction.

In order to overcome this hindrance, this paper proposes a modified Gabor filtering technique based on NB classifier (MGF + ENBA) for overcoming the issue. The MGF technique extracts only the most important features from the available data in an efficient manner. The ENBA algorithm is then used for the classification of facial expression in a more accurate way with the emotion detection such as sad, happy, angry, fear, and neutral emotional features. The rest of the paper is organized in the following manner. Section 2 provides the gist of important works in the literature in the emotion detection, Sect. 3 explains the proposed method in detail, Sect. 4 elaborates about the experimental result and discussion. The conclusion is presented in Sect. 5.

2 Related Work

Kim et al. [17] gave special attention for the utilization of contrast representation which can embed a distinct feature factor for a discriminative intend. The contrast representation that embed a deep neural network is used for the comparison of the given image with that of a reference. The authors attempted to use a generic reference which are estimated from the given image. As a consequence, the DNN is deployed which can combine to form an iterative model that has an end-end training. In the proposed network, an attempt to disintegrate the expressive factor in a couple of steps which includes the learning development framework is carried out. The experiments were carried out using public datasets available and it is seen that the propose method has high detection accuracy.

Dewi et al. [18] performed a critical review on the introduction of facial emotions and developed the LGBP-TOP combined with SVM which tend to focus on the images of faces and lips alone. A novel algorithm was developed for extracting the features of lips and eyes using the 3D blocks which are then used as the input for LGBP-TOP technique. Later, the results are classified by using the SVM technique. The experimental results proved an overall accuracy of 82% and found to perform better than the entire face and also specific areas such as eyes and lips.

Luo et al. [19] presented the PCA method which is based on statistical analysis that is extracted from the gray scales of the full image. The gray scaled features are highly sensitive toward the environment, and hence, a hybrid method that combines the PCA and LBP was proposed. The LBP is determined to extract the locally available gray features on the regions surrounding the month and eyes. The SVM methods are used for the detection of facial emotions. The experimental results prove that the proposed method can able to classify the various expressions in a more effective manner and can get more rate of recognition than the present traditional methods.

Krestinkaya and James [20] used the simplest method for the feature selection which is also effective. The authors have proposed an straight forward technique for the recognition of emotions for reducing the issue of the inter-class mismatch among the pixels during the process of classification. The proposed method uses the application of normalizing the pixels for the removal of offsets cause due to intensity

followed by the Min–Max metrics in an nearest classifier that are competent for the outlier suppression of features which forms the outliers. The experimental results indicate that there is a significant improvement in the performance of 18% when the min_max classification process is tested on the JAFFE base. The proposed technique outperforms all the other existing methods that are based on template matching.

3 Proposed Methodology

The main objective of the proposed modified Gabor filter with NB algorithm (MGF + ENBA) is to improve the FER performance.

3.1 Image Dataset

All the facial emotion classification sets that are used in the proposed method are from couple of databases, the JAFFE [21], and Cohn-Kanade [22]. The former data base contains 239 images of 7 basic facial expressions. These images are taken from the female models belonging to japan. The emotions expressed in each of the picture are tested for its subjectivity on 60 different volunteers. The sample set of images is shown in Fig. 1.

3.2 Image Acquisition

The images that are used in the proposed method are static in nature. In order to capture the images of various expressions, a Panasonic model of camera with a focal length of 7 mm is made used. The JPEG format is used for the images with 24-bit coloring. The resolution was fixed as 4320 * 3240 pixels. The maximum distance between the camera and the person was kept as 4 feet and the images of seven basic expressions of each of the persons are taken. The image sequence is shown as in Fig. 2.

3.3 Image Preprocessing

The preprocessing is an important step in the image processing and not exemption in case of facial expression detection. The aim of preprocessing is to pick out the images that have normal intensity with uniform shapes and size and are represented only the face that has emotion. The preprocessing helps in reducing the effect of illumination and as well as lighting. The representation of the expression can be difficult for

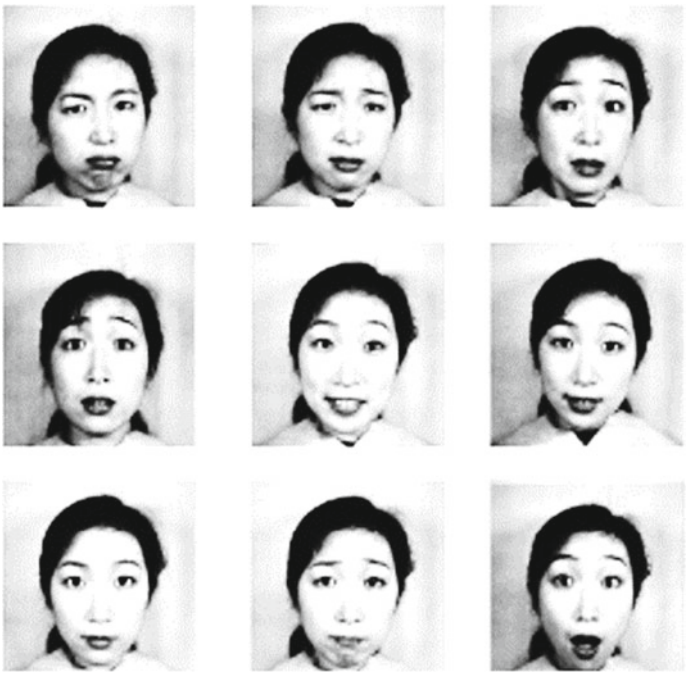


Fig. 1 Sample static images from JAFFE database

Fig. 2 Image sequences from Cohn-Kanade database

translating and rotation. To overcome these hindrances of the pointless changes, the facial images are geometrically processed before entering into classification. The preprocessing technique becomes more important in the FER task. The objective of preprocessing is to gather images that have normal intensity and uniform size and shape and depicts only the face expressions.

The preprocessing is also expected to remove the effect of lighting. The area pertaining to the face of a given image is detected using the Viola-Jones method using the AdaBoost algorithm. The method uses an object detection technique that provides a comprehensive detection of objects in a real-time. It is initially designed for the issue of face detection. The significant features that are used by the viola are obtained from the pixels that are selected from a rectangular area which is imposed over the picture and shows high level of sensitivity for the vertical and the horizontal lines. AdaBoost has an adaptive learning mechanism which can be used as a conjunction with all other learning methods for performance improvement. AdaBoost is considered as adaptive as the subsequent classifiers are built iteratively and are made to fix the instances that are misclassified by the former classifiers. At every iteration, a distributed weight is updated in such a way that the weights of every incorrect classification are increases which enables the new classification to be focused.

The last step in the preprocessing phase is the detection of facial picture that has emotion within the max intensity level of a given emotion. The mutual information-based method is used for the face detection. For every frame, this information presents in-between the present frame and the initial one is computed and the particular frame which has the mutual kind of information is then selected as the one which represents a particular emotion with maximum level of intensity. At last, the images are brought back to the initial size. Figure 3 shows images after the preprocessing.

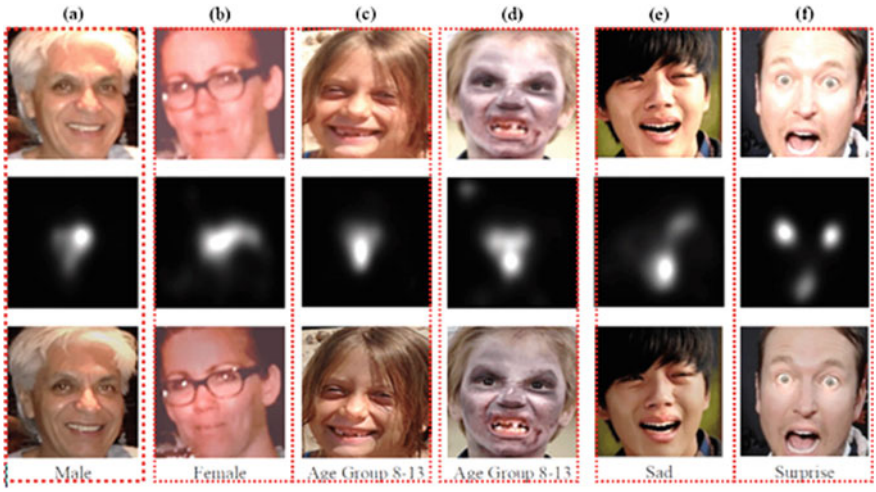


Fig. 3 Images after preprocessing step

3.4 Feature Extraction Using Modified Gabor Filters (MGFs)

The Gabor filters (GFs) are generally considered as one among the best choice for getting a local frequency information. However, they commonly suffer from couple of important limitations. The total bandwidth in case of a GF is limited to a single octave and also these filters are not optimal if one of them seeks spectral information that has more localization. Log-based GFs help to overcome this limitation. The DC component is nullified here and can be built with an approximate bandwidth that can be able to produce the filter that has minimal spatial extends. Log-based filters in the frequency issue are defined in terms of polar bi-coordinates using $H(f, \theta) = Hf \times H\theta$, where Hf depicts the radial item and $H\theta$, the angular component.

$$H(f,\theta) = \exp\left\{\frac{-\left[\ln\left(\frac{f}{f_0}\right)\right]^2}{2\left[\ln\left(\frac{\sigma_f}{f_0}\right)\right]^2}\right\}\exp\left\{\frac{-(\theta-\theta_0)^2}{2\sigma_\theta^2}\right\} \tag{1}$$

Here, the $f()$ is the frequency at the center of the filter and θ_0 represents the direction of the filter. Σf is the constant value that defines the bandwidth (B) of the octaves. The co-efficient of the MGF is then calculated in an off-line speedup. The GF filters are with different orientations and are employed through the co-efficient that are calculated, respectively. This gives a variety of filters (Figs. 4 and 5).

Fig. 4 Eye images used for training and testing

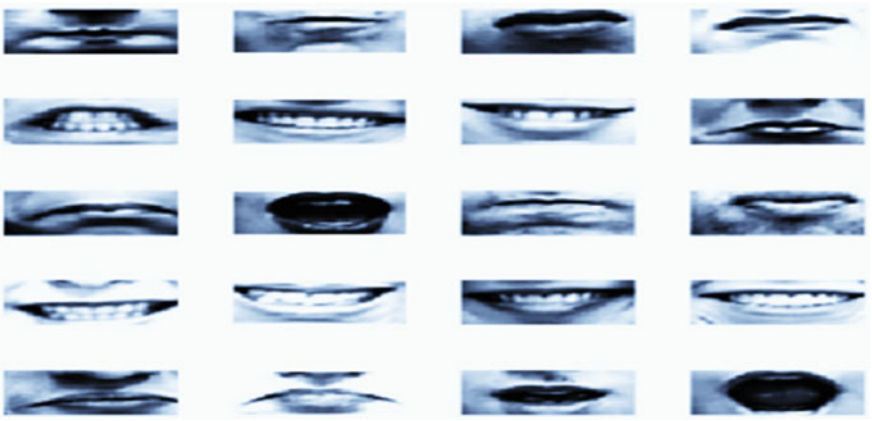


Fig. 5 Mouth images used for training and testing

3.5 FER Classification Using Enhanced Naïve Bayes Algorithm (ENBA)

The NB classifier is based on the probability that has displayed effectiveness in many problems of classification [7]. The NB classifiers are simple but are effective in classification which are based on Bayes theorem. The Navie Bayes theorem is trained by using the images that are captured in a supervised learning environment. It is assumed that the present of a given feature in a class is not related with the presence of any other feature. If c represents the class variable, then the features are represented using the set $\{f_1, f_2, \ldots, f_k\}$, the decision of the classifier is made by using the formula

$$C = \arg\max_{c} \left\{ p(c) \prod_{j=1}^{k} p(f_i|c) \right\} \tag{2}$$

where the $p(c)$ represents the total samples in the set $\{c\}$/the total no. of samples are the conditional table which are learnt in the example and k represents the feature vector length. Although there is an existence of independent assumptions, the NB has proven to be a good choice for classification for real-time data in par with many complex classifiers. The overall architecture of the proposed scheme is given in Fig. 6.

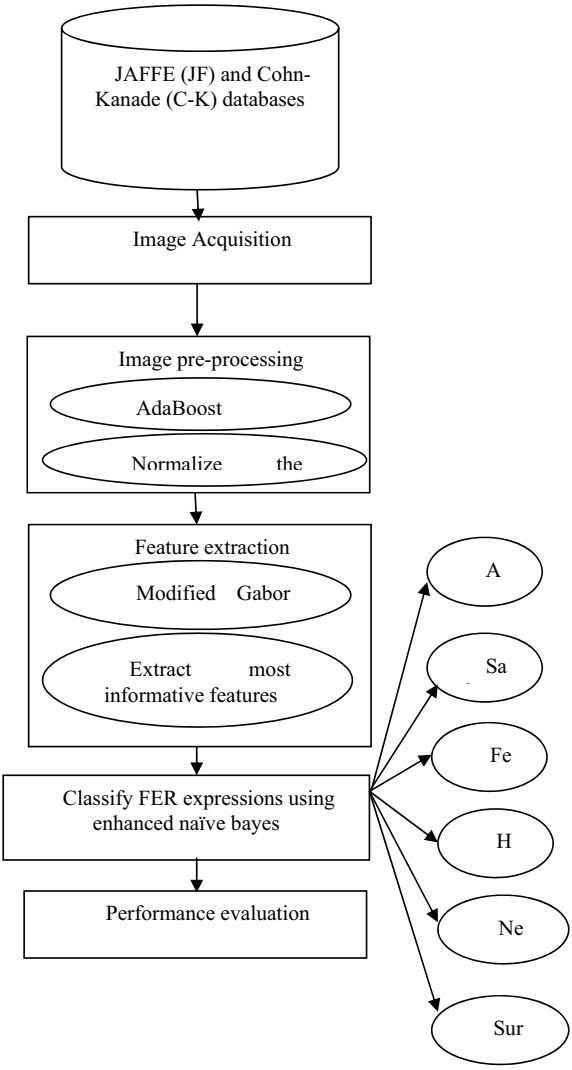


Fig. 6 Overall block diagram of the proposed system

4 Experimental Result

The JAFFE repository has a total of 596 images' sequence ranging to 124 subjects in the age range between 18 and 35. The DB includes faces that show the basic 7 emotions for which there were {45, 98, 56, 58, 78, 35, 142} images, respectively. The proposed method is compared with that of the nearest neighbor classification and SVM for evaluating the JAFFE database along the MMSNN algorithm. The

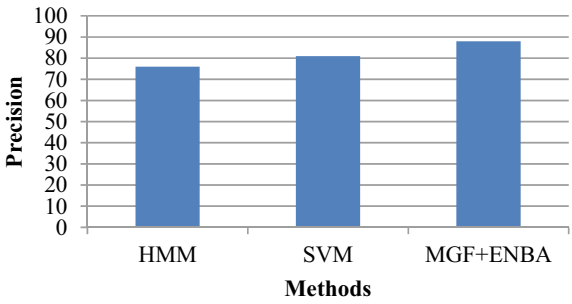


Fig. 7 Precision comparison

performance metrics chosen are precision, recall, and F-measure for as the measure of accuracy.

Precision

The precision can be viewed as the computational parameter of accuracy or a quality and is calculated using the formula,

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \tag{3}$$

The recall is the measure of the completeness or in other words, it is a quality parameter. Usually, the high value of precision indicates that the algorithm has given back more relevant results. In the classification problem, the precision of a given class is the total count of true positives divided by the total elements that belong to positive class.

From Fig. 7, it is observed that the plot is drawn taking the existing and the proposed method and compared for its precision. In the X axis, the methods are assigned and in the Y axis, the values of precision are plotted. It is observed that the proposed method has more precision than the other popular methods such as the HMM and SVM. This conclude that the proposed method can increase the facial expressions.

Recall

Recall is generally defined as the no. of appropriate documents received by a given search divided by the total count of existing such documents and is calculated as

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \tag{4}$$

From Fig. 8, it is seen that the comparative metric is evaluated between the proposed and the existing method. In the X axis, the methods are assigned and in the Y axis, the values of recall are plotted. It is observed that the proposed method has more precision than the other popular methods such as the HMM and SVM. This conclude that the proposed method MGF + ENBA can increase the detection

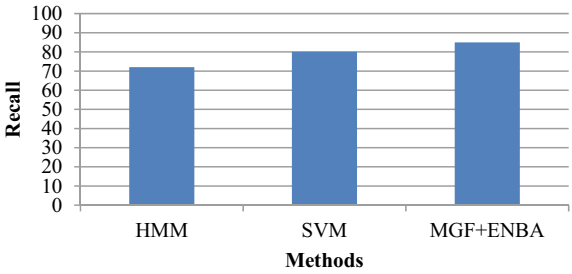


Fig. 8 Recall

performance of the facial expressions such as sad, happy, anger, fear, neutral, and surprise expressions accurately for the given databases.

F-Measure

F_1-score calculated as

$$F_1\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{5}$$

From Fig. 9, it is seen that the comparative metric is evaluated between the proposed and the existing method. In the *X* axis, the methods are assigned and in the *Y* axis, the values of *F*-Measure are plotted. It is observed that the proposed method has more precision than the other popular methods such as the HMM and SVM. This conclude that the proposed method MGF + ENBA can increase the detection performance of the facial expressions such as sad, happy, anger, fear, neutral, and surprise expressions accurately for the given databases.

Accuracy

Accuracy is the correctness of a given model and it is calculated as the total count of actual parameters ($T_p + T_n$) of classification that are segregated by the sum of all of the classification's parameter ($T_p + T_n + F_p + F_n$), the same is computed as

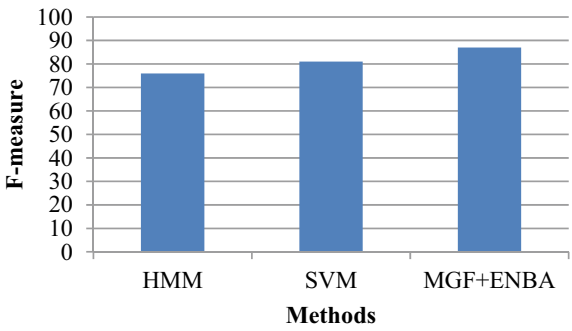


Fig. 9 *F*-measure

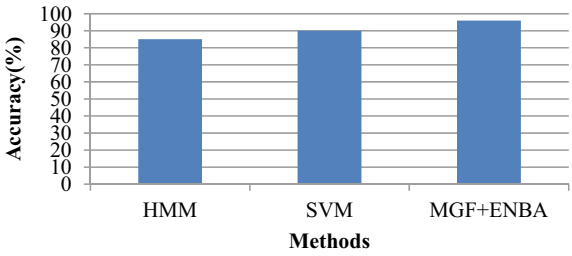


Fig. 10 Accuracy

$$\text{Accuracy} = \frac{T_\text{p} + T_\text{n}}{\left(T_\text{p} + T_\text{n} + F_\text{p} + F_\text{n}\right)} \tag{6}$$

From Fig. 10, it is seen that the comparative metric is evaluated between the proposed and the existing method. In the X axis, the methods are assigned and in the Y axis, the values of accuracy are plotted. It is observed that the proposed method has more precision than the other popular methods such as the HMM and SVM. This conclude that the proposed method MGF + ENBA can increase the detection performance of the facial expressions such as sad, happy, anger, fear, neutral, and surprise expressions accurately for the given databases.

5 Conclusion

In this paper, an efficient method for the calculation of accurate FER rate using the MGF + ENBA technique. The FERs are based on the hybrid facial regions which are taken for investigation. The part of the face that shows the most discriminative information on the facial emotional expression classification. The vector representing the features are extracted from original image. Then, the significant features are alone selected based on the mutual information. These are classified using the ENBA classifier. The experiments are carried out using the JAFFE database images and it is found that the proposed MGF + ENBA algorithm gives more accuracy and better performance than other competitive methods such as HMM and SVM algorithms.

References

1. Zeng N et al (2018) Facial expression recognition via learning deep sparse autoencoders. Neurocomputing 273:643–649
2. Yang H, Ciftci U, Yin L (2018) Facial expression recognition by de-expression residue learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition
3. Kołakowska A, Landowska A, Szwoch M, Szwoch W, Wrobel MR (2014) Emotion recognition and its applications. In: Human-computer systems interaction: backgrounds and applications, Chap 3. Springer International Publishing, Cham, pp 51–62

4. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. NIPS 1:4
5. Liu P, Han S, Meng Z, Tong Y (2014) Facial expression recognition via a boosted deep belief network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1805–1812
6. Zhao G, Pietikainen M (2007) Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans Pattern Anal Mach Intell 29(6):915–928
7. Azcarate A et al (2005) Automatic facial emotion recognition. Universiteit van Amsterdam 1–6
8. Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B, Cukierski W, Tang Y, Thaler D, Lee D-H et al (2013) Challenges in representation learning: a report on three machine learning contests. In: International conference on neural information processing. Springer, Berlin, pp 117–124
9. Dhall A, Ramana Murthy O, Goecke R, Joshi J, Gedeon T (2015) Video and image based emotion recognition challenges in the wild: Emotiw 2015. In: Proceedings of the 2015 ACM on international conference on multimodal interaction. ACM, pp 423–426
10. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
11. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556
12. Agrawal R, Jalal AS, Arya KV (2019) Fake fingerprint liveness detection based on micro and macro features. Int J Biometrics 11(2):177–206
13. Chawla D, Trivedi MC (2019) Proposed a new multidimensional face recognition for surgically altered faces in security surveillance. Int J. Eng Adv Technol (IJEAT) 8(6S):1021–1026. ISSN: 2249-8958
14. Agarwal R, Jalal AS, Arya KV (2020) A multimodal liveness detection using statistical texture features and spatial analysis. Multimed Tools Appl 79:13621–13645
15. Rifai S, Vincent P, Muller X, Glorot X, Bengio Y (2011) Contractive auto-encoders: explicit invariance during feature extraction. In: Proceedings of the 28th international conference on international conference on machine learning. Omnipress, pp 833–840
16. Masci J, Meier U, Cireşan D, Schmidhuber J (2011) Stacked convolutional auto-encoders for hierarchical feature extraction. In: International conference on artificial neural networks. Springer, Berlin, pp 52–59
17. Kim Y et al (2017) Deep generative-contrastive networks for facial expression recognition. arXiv preprint arXiv:1703.07140
18. Dewi RRK, Sthevanie F, Arifianto A (2019) Face expression recognition using local Gabor binary pattern three orthogonal planes (LGBP-TOP) and support vector machine (SVM) method. J Phys Conf Ser 1192(1)
19. Luo Y, Wu C, Zhang Y (2013) Facial expression recognition based on fusion feature of PCA and LBP with SVM. Optik-Int J Light Electron Opt 124(17):2767–2770
20. Krestinskaya O, James AP (2017) Facial emotion recognition using min-max similarity classifier. In: 2017 International conference on advances in computing, communications and informatics (ICACCI). IEEE
21. Kanade T, Cohn JF, Tian Y (2000) Comprehensive database for facial expression analysis. In: Proceedings of the fourth IEEE international conference on automatic face and gesture recognition, Grenoble, France, pp 46–53
22. Chawla D, Trivedi MC (2018) A review paper on face recognition for security surveillance of surgically altered faces. J Adv Res Dyn Control Syst 10(12-special issue)

A Deep Neural Network (DNN) Approach for Recommendation Systems

Shashi Shekhar, Anshy Singh, and Avadhesh Kumar Gupta

Abstract In different commercial platforms, Recommendation System (RS) is widely used for providing recommendations to users. In various areas, RS is utilized broadly and in E-Commerce sites, product recommendations are discovered during an active user interaction by RS. In recent decades, some key challenges are faced due to tremendous growth in user as well as products. Moreover, in RS, computation of right product and active user is a major task. User inclination and socio-demographic behavior are considered in existing works for recommending a product. In recommendation systems, one of the major algorithm used is Collaborative Filtering (CF) algorithms. This algorithm is simple as well as effective. However, further enhancement of recommendation result's quality is limited by data sparsity and scalability of this technique. Previous technique's problems are addressed effectively in proposed technique and user preference on balance feature analysis and products are evaluated. Therefore, proposed a model using the combination of deep learning technology and CF recommendation algorithm with three major stages, namely, preprocessing, representation of features and rating score prediction using DNN. At first, from log files, redundant and unnecessary data are removed in preprocessing module. There is an unwanted files like repeated tags, repeated similar products, removing invalid values, last visit and elapsed time. In feature representation stage, Quadric Polynomial Regression—QPR-based feature representation technique is used. The traditional matrix factorization algorithm is enhanced for obtaining accurate latent features. At last, DNN model is fed using these latent features as input data, which is a second stage of proposed model. Rating scores are predicted using this. From Amazon dataset, user data based on behavior is obtained and used in experimentation. There are 18,501 product reviews in Amazon product dataset. From Amazon web

S. Shekhar (✉) · A. Singh
Department of Computer Engineering and Application, GLA University, Mathura, India
e-mail: shashi.shekhar@gla.ac.in

A. Singh
e-mail: anshy.singh@gla.ac.in

A. K. Gupta
USCI, Karnavati University, Gujrat, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_37

services, collected the dataset information that joins with administrative services. Based on metrics like $F1$-measure, Recall (R) and Precision (P) proposed Deep Neural Network (DNN) method is evaluated experimentally and highest value of those metrics are produced when compared with state-of-the art techniques like K-Nearest Neighbor (K-NN), Artificial Neural Networks (ANN).

Keywords Recommendation system (RS) · Collaborative filtering (CF) · Quadric polynomial regression (QPR) · Deep neural network (DNN)

1 Introduction

In daily life, more intelligent products are being applied increasingly due to artificial intelligence technology's development. For people in various aspects, convenience is provided by this. From massive Internet data, valuable information are provided to users effectively using personalized recommendation system's intelligent recommendation function. So, in various network platforms like shopping, music and movie, it is used widely. Online commercial centers make their very own benefit dependent on their notices while business metric has business enthusiasm for ranking higher on suggestions to draw in more users [1]. Only popular products are presented in tradition retail, while variety of products are presented in online. It is highly difficult for users for comprehending the accuracy of items given by recommender frameworks, because of tremendous web information [2]. From related dataset, user interest are extracted using recommender system and for big dataset, quality recommendations are provided using this. In E-commerce, recommender system is a significant part, data mining and machine learning techniques are used in this [3] for filtering unseen information for predicting items of user interest. A special recommender system type is an intelligent system, used for exploiting historical user ratings [4–9]. In order to engage users and satisfying them with personalized recommendations in information explosion age, Recommendation systems are vital. In modern social media, entertainment and E-commerce platform, personalized content are expected by users. But, available model capacity and user-item interactions, limits the recommendations effectiveness.

The sparsity problem can be alleviated by leveraging the ability of higher order reasoning. From past interactions like purchases, ratings and clicks, relevance between items and users are established in Collaborative Filtering (CF), where consumption of similar items by similar users are assumed. It is a successful and popular technique [10, 11].

In a recommendation system, highly important part is recommendation algorithm and recommendation result's quality and system's performance are defined directly by this [12]. There are two classes of commonly utilized algorithms. They are, collaborative filtering techniques and content-based techniques. Through the extra information like items attributes, user profiles, document content, user and item's portraits are constructed in content-based techniques for making recommendations [13].

In many situation, information utilized for constructing portraits are difficult to obtain or sometimes, it may be fake one. From significant limitations, its application and performance are suffered.

In recommendation systems, most widely used algorithm is Collaborative filtering algorithms [14, 15]. Information regarding items or user are not required in this algorithm, which makes the difference from content-based techniques. Based on interaction information between items and users like rating, browsing and clicks, they make recommendations accurately. It is an effective as well as simple technique. Algorithm performance is limited by data sparsity and Internet's rapid development. So, for enhancing recommendation performance, some other techniques are developed by researchers [16, 17].

In various research areas like machine translation, learning programs, question answering, computer vision and many other domains, state of art performance are obtained using deep learning. Over traditional linear models, complex nonlinear transformation's noticeable advantages are demonstrated by successful integration techniques in recommendation systems [18, 19]. In various fields like language processing, speech recognition, computer vision, great success is achieved by Deep Neural Networks (DNNs) in recent years. However, on recommendation systems, there exist some studies with these technologies.

Following are the steps of proposed work:

1. From log files, redundant and unnecessary data are removed using preprocessing module. There are some unwanted files like repeated tags, repeated similar products, removing invalid values, elapsed time since last visit.
2. Traditional matrix factorization algorithm is improved to obtain highly accurate latent features in QPR-based feature representation module in second stage. Items and user features are obtained using user-item rating matrix, which is elaborated in Sect. 3.
3. Then, in proposed model's second part, DNN model is fed with these latent features as input data and rating scores are predicted using this. Probability values representing scores probability are obtained in output layer and scores having highest probability is used as the result of prediction.

The rest of paper is organized as follows: proposed methodology on RS system using deep learning algorithm in Sect. 2. In Sect. 3, results and discussion of this study are presented. In Sect. 4, conclusions and future work are provided.

2 Proposed Methodology

For assessing gathered information, over dataset, applied the proposed DNN as input and for framing sorted out data, explicit feedback are processed. Data having user information attributes are integrated with user behavior ranking of various products. In original database, all information are valid, so they are extracted and pre-processed the large amount of significant data.

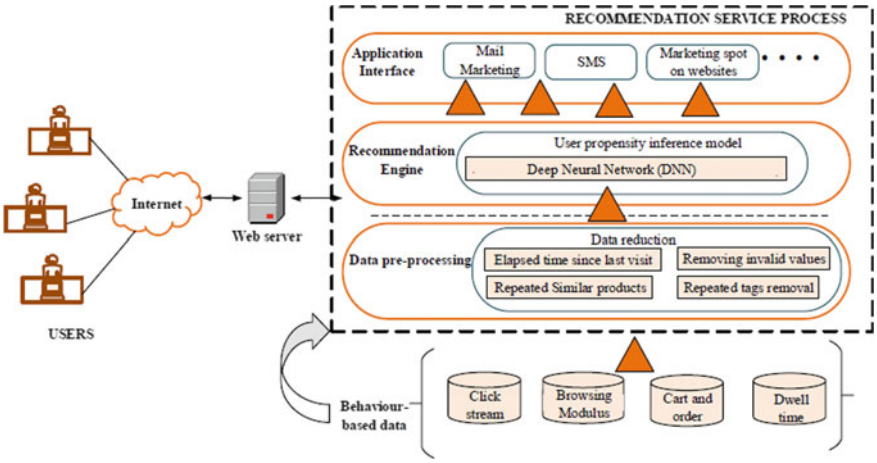


Fig. 1 The proposed recommendation model

From log files, redundant and unnecessary data are removed using preprocessing module. There are some unwanted files like repeated tags, repeated similar products, removing invalid values, elapsed time since last visit. Traditional matrix factorization algorithm is improved to obtain highly accurate latent features in QPR-based feature representation module in second stage. At last, DNN model is given with these latent features as input, which is a proposed model's second part and rating scores are predicted using this. Figure 1 shows the proposed recommendation model.

2.1 Recommendation Service Process

Three main phases are included in this process, namely, Recommendation Engine, Data Preprocessing and Data Collection.

2.1.1 Data Collection

From Amazon dataset, obtained user's data based on behavior in this component [20]. There are product list file and user's behavioral file in principle records of collected dataset. Personal background like user's browsing modules, dwell time, click stream are there in user's behavioral file. From E-commerce site, collected data type includes, time category, product ID, clicks count, purchases, session ID, product name like electronics, dresses, books, etc. Within the user requests, collected data is repeatedly scheduled according to user data.

2.1.2 Data Preprocessing

From log files, redundant and unnecessary data are removed using preprocessing module. There are some unwanted files like repeated tags, repeated similar products, removing invalid values, elapsed time since last visit. In general, this work is performed using rating matrix, where user feedback are represented as ratings. In behavior database, user behavior matrix is used for various behavior types for serving as an every user's implicit ratings of various products.

This matrix's scarcity (S) is computed as,

$$S = \frac{N_{\text{non-zero}} * N}{N_{\text{total}}}$$

where non-zero values count is represented as $N_{\text{non-zero}}$, recommendation system's required value is represented as N and in matrix, total count values are represented as N_{total}.

2.1.3 Recommendation Engine

See Fig. 2.

2.2 Feature Representation

Based on rating matrix, items and user's features needs to be obtained as shown in Fig. 2. Assume $\in \mathbb{R}^{n \times m}$ as n user's user-item rating matrix to m items, jth item's ith user's rating score is represented as R_{ij}, $R_{ij} = 0$, if there is no rating record. Moreover, assume $U \in \mathbb{R}^{n \times a}$, as latent feature matrix of user, ith user's features are represented using ith row of vector U_i and features dimension are represented as a. Ratings as Features (RaF): A type of feature representation technique is RaF. In this, user's rating data are regarded directly as user feature. In specific, $U = R$ and items features are $V = R^T$. An effective and simple technique is RaF, which has some drawbacks too. Highly sparse nature is exhibited by rating matrix. There exist a huge amount of missing data.

There will be a uniqueness among user ID or item. So, it is assumed as a required feature. As, ID is a categorical variable, it cannot be summed or compared or used with other mathematical operations.

$$U_i = \text{NN}(\text{OneHot}(i)),$$

$$V_j = \text{NN}(\text{OneHot}(j)),$$

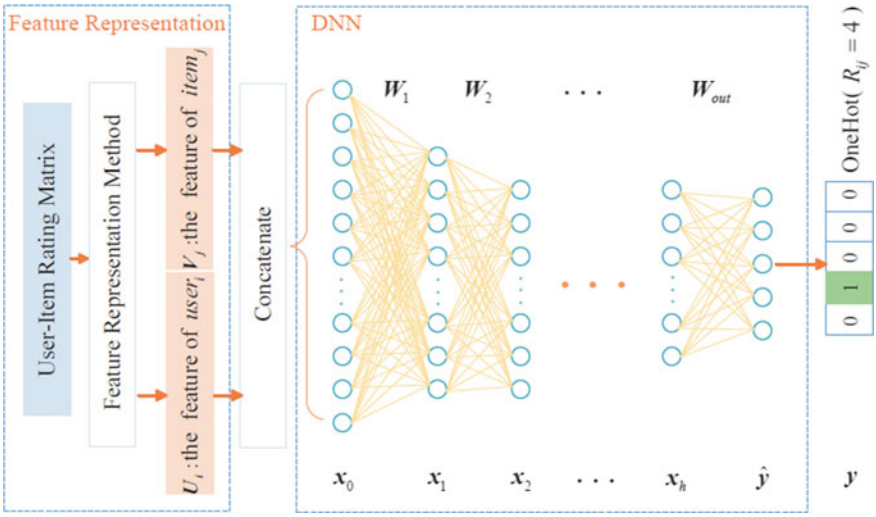


Fig. 2 The framework of DNN model

where usage of onehot technique for encoding IDi is represented as OneHot(i), with a specified dimension, a zero vector is generated by this and value of 1 is assigned to vector's ith position. With x as input, neural network's output is represented as NN(x).

Singular Value Decomposition (SVD): With decreased feature representations and dimensions, a strong performance can be achieved using matrix factorization technique. So, from a matrix, features can be learnt easily using this.

The SVD is expressed as follows:

$$R = U.S.V^T$$

where $S \in \mathbb{R}^{n \times m}$ represents diagonal matrix and matrix R's eigenvalues are represented in main diagonal elements.

Quadric Polynomial Regression (QPR): QPR-based new feature representation technique is proposed in this paper by considering above-listed NCF and RaF's disadvantages. Inaccurate results produced due to missing values preprocessing is avoided in this technique and between features, correlation are considered.

For feature vector x in traditional Quadric Polynomial Regression model, following expression is used for fitting corresponding supervised value y.

$$\hat{y} = z + \sum_{i=1}^{l} w_i x_i + \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} W_{ij} x_i x_j$$

where vector x's dimension is represented using a parameter l, constant term coefficient is represented as Z, first-order coefficients are represented as w, second order

coefficients are represented as $W \in \mathbb{R}^{1\times1}$. So, in proposed technique, there are $x = (U_u, V_v)$and$y = R_{uv}$.

$$\hat{R}_{uv} = z + \sum_{i=1}^{a} w_i U_{ui} + \sum_{j=1}^{b} w_{j+a} V_{vj} + \sum_{i=1}^{a-1} \sum_{j=i+1}^{a} W_{ij} U_{ui} U_{uj} + \sum_{i=1}^{b-1} \sum_{j=i+1}^{b} W_{i+a,j+a} V_{vi} V_{vj} + \sum_{i=1}^{a} \sum_{j=1}^{b} W_{i,j+a} U_{ui} V_{vj}$$

$$p_u = \sum_{i=1}^{a} w_i U_{ui} + \sum_{i=1}^{a-1} \sum_{j=i+1}^{a} W_{ij} U_{ui} U_{uj}$$

$$q_v = \sum_{j=1}^{b} w_{j+a} V_{vj} + \sum_{i=1}^{b-1} \sum_{j=i+1}^{b} W_{i+a,j+a} V_{ui} V_{vj}$$

2.3 Deep Neural Network (DNN)

Figure 2 shows the model of DNN, where items and user's latent features are given as inputs and rating scores are predicted using forward propagation algorithm [21]. Items and user's latent features are used for concatenating input vector x_0 and input layer based on this model. So, for any record R_{ij},

$$x_0 = \text{concatenate}\left(U_i, V_j\right)$$

where two vectors are concatenated using a function called concatenate(). Following expression is used for obtaining first hidden layer output during x_0 passing through it.

$$x_1 = \text{activation}(W_1 x_0 + b_1)$$

where between first hidden layer and input layer, weight matrix is represented as W_1, bias vector is represented as b_1, activation function is indicated as activation(), nonlinear neural network model are made by designing this and multilayer neural networks have added advantages.

In DNN model, ReLU, tanh and sigmoid functions are included as an activation function.

$$x_1 = \text{Relu}(W_l x_{l-1} + b_l)$$

Prediction of user's rating score R_{ij} is training goal in output layer. Supervised value $y = \text{OneHot}(R_{ij})$ is obtained using a hot encoding technique. So, for obtaining y's corresponding position's prediction value, output results are transformed using softmax technique and it is given by,

$$\hat{y} = \text{softmax}(W_{\text{out}} x_h + b_{\text{out}})$$

where hidden layers are represented as h, last hidden layer's output is represented as x_h, output layer's weight value is represented as W_{out} and bias value is represented as b_{out}.

2.4 Training Model

At first, features are obtained using QPR. Following gives every parameter's updating rule.

$$z = z - \eta \sum_{u,v} \Delta_{uv}$$

$$p_u = p_u - \eta \sum_{v} \Delta_{uv}$$

$$q_v = q_v - \eta \sum_{u} \Delta_{uv}$$

$$W_{ij} = W_{ij} - \eta \sum_{u,v} \Delta_{uv} U_{ui} V_{vj}$$

$$U_{ui} = U_{ui} - \eta \sum_{v} \left(\Delta_{uv} \sum_{j=1}^{b} W_{ij} V_{vj} \right)$$

$$V_{vj} = V_{vj} - \eta \sum_{u} \left(\Delta_{uv} \sum_{i=1}^{a} W_{ij} U_{ui} \right)$$

Bias vector b and weight matrix W's learning form base for DNN training. Between ith neuron in lth layer and jth neuron in $(l-1)$th layer, connection weight is represented by defining $W_{l,i,j}$. In specific, input layer represents $0th$ layer. In lth layer, on ith neuron bias value is represented as $b_{l,i}$.

3 Results and Discussion

The proposed RS based on DNN's estimated results, metrics and dataset is explained in this. With respect to metrics like Recall, $F1$ measure and Precision, proposed techniques effectiveness is estimated. Baseline models like K-NN and ANN are used for making comparison. Amazon product dataset is used in DNN.

There are 18,501 product reviews about category, brand name, item ID and item in Amazon product dataset [20] and subfields like response number, view, pct purchase, bytes, change quantity, add to cart and action type are included in purchase's metadata information.

Also referenced the information like URL address, Timestamp, Rstat, Rbytes, Session ID, Server IP and Client IP. From Amazon web services, collected the dataset information that joins with administrative services. Proposed DNN approach produces better results when compared with K-NN and ANN as revealed in experimentation results.

Precision (P): It is defined as a fraction of total recommended items count from relevant recommended items for target user. This measures for computing values whether higher values specify better performance and it is given by,

$$P = \frac{n_{\text{rs}}}{n_s}$$

Recall (R): It states that proportion of all relevant recommended items from various products for target user. As like precision, recall measures for computing values whether higher values specify better performance and it is given by,

$$R = \frac{n_{\text{rs}}}{n_s}$$

where products count favored by target user is represented as n_r.

***F*1-measure:** It is defined by precision and recall, it is expressed as (Table 1),

$$F1 = \frac{2 * P * R}{P + R}$$

Figure 3 shows the precision value comparison of three classification techniques. Amazon dataset is used for achieving precision value. Increase in neighborhood

Table 1 Performance comparison metrics versus various recommendation systems

Methods	Precision (P) (%)	Metrics recall (R) (%)	$F1$-measure (%)
ANN	80	78	75
K-NN	85	83	78
DNN	87	85	82

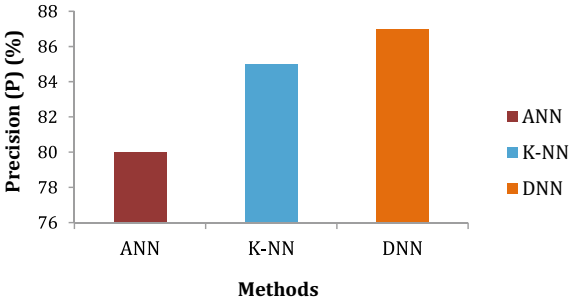


Fig. 3 The comparison of precision (p) measures for various methods

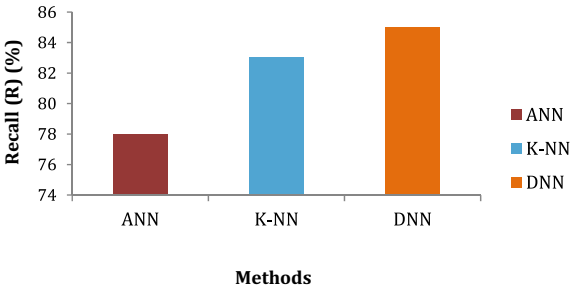


Fig. 4 The comparison of recall (r) measures for various methods

increases DNN techniques precision value. When compared with traditional techniques like K-NN and ANN, better precision values are produced by DNN. Around 87% of precision results are produced by proposed DNN as indicated in that figure and it is a higher value when compared with 80% precision results of ANN and 85% precision results of K-NN.

Figure 4 shows the recall value comparison of three classification techniques. Amazon dataset is used for achieving recall value. Increase in neighborhood increases DNN techniques recall value. When compared with traditional techniques like K-NN and ANN, better recall values are produced by DNN. Around 85% of recall results are produced by proposed DNN as indicated in that figure and it is a higher value when compared with 78% recall results of ANN and 83% recall results of K-NN.

Figure 5 shows the $F1$-measure comparison of three classification techniques. Amazon dataset is used for achieving $F1$-measure value. Increase in $F1$ value increases size of neighborhood. Increase in proposed DNN's $F1$ value increases recommendation list with neighborhood size. Around 82% of $F1$-measure results are produced by proposed DNN as indicated in that figure and it is a higher value when compared with 75% $F1$ measure results of ANN and 78% $F1$ measure results of K-NN.

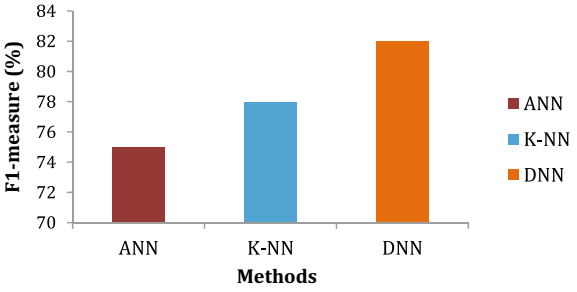


Fig. 5 The comparison of $F1$-measure for various methods

4 Conclusion and Future Work

User data behaviors like views count, purchased product, clicks and purchase histories are used for proposing DNN model. Therefore, proposed a model using the combination of deep learning technology and CF recommendation algorithm with three major stages, namely, preprocessing, representation of features and rating score prediction using DNN. At first, from log files, redundant and unnecessary data are removed. In feature representation stage, QPR-based feature representation technique is used. The traditional matrix factorization algorithm is enhanced for obtaining accurate latent features. At last, DNN model is fed using these latent features as input data, which is a second stage of proposed model. Rating scores are predicted using this. User behavior matrix is constructed in proposed technique and purchased product's frequency is pre-processed for predicting users via feature vectors of every product and neighbor. From Amazon dataset, user data based on behavior is obtained and used in experimentation. Based on metrics like $F1$-measure, Recall (R) and Precision (P) proposed method is evaluated experimentally and highest value of those metrics are produced when compared with state-of-the art techniques. In RS, multi-dimensionality concept should be incorporated in future.

References

1. Davoudi A, Chatterjee M (2018) Social trust model for rating prediction in recommender systems: effects of similarity, centrality, and social ties. Online Soc Netw Media 7:1–11
2. Sinha BB, Dhanalakshmi R (2019) Evolution of recommender system over the time. Soft Comput 23:12169–12188
3. Christudas BCL, Kirubakaran E, Thangaiah PRJ (2018) An evolutionary approach for personalization of content delivery in e-learning systems based on learner behavior forcing compatibility of learning materials. Telematics Inform 35(3):520–533
4. Tarus JK, Niu Z, Mustafa G (2018) Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning. Artif Intell Rev 50(1):21–48
5. CireşAn D, Meier U, Masci J, Schmidhuber J (2012) Multi-column deep neural network for traffic sign classification. Neural Netw 32:333–338

6. Shekhar S, Mohan N (2021) Sentiment classification using hybrid bayes theorem support vector machine over social network. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
7. Shekhar S, Varshney N (2021) Hybridization of Social Spider Optimization (SSO) Algorithm with Differential Evolution (DE) using super-resolution reconstruction of video images. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
8. Shekhar S, Sharma DK (2020) Computational intelligence for temporal expression retrieval in code-mixed text. In: 2020 international conference on power electronics & IoT applications in renewable energy and its control (PARC). Mathura, Uttar Pradesh, India, pp 386–390
9. Shekhar S, Sharma DK, Sufyan Beg MM (2020) Language identification framework in code-mixed social media text based on quantum LSTM—the word belongs to which language? Modern Phys Lett B 34(6):2050086
10. Richardson F, Reynolds D, Dehak N (2015) Deep neural network approaches to speaker and language recognition. IEEE Sig Proc Lett 22(10):1671–1675
11. Liu J, Jiang Y, Li Z, Zhang X, Lu H (2016) Domain-sensitive recommendation with user-item subgroup analysis. IEEE Trans Know Data Eng 28(4):939–950
12. Hwangbo H, Kim YS, Cha KJ (2018) Recommendation system development for fashion retail e-commerce. Electron Commer Res Appl 28:94–101
13. Wei J, He J, Chen K, Zhou Y, Tang Z (2017) Collaborative filtering and deep learning based recommendation system for cold start items. Exp Syst Appl 69:29–39
14. Ebesu T, Shen B, Fang Y (2018) Collaborative memory network for recommendation systems. In: The 41st international ACM SIGIR conference on research and development in information retrieval, pp 515–524
15. Nilashi M, Ibrahim O, Bagherifard K (2018) A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques. Exp Syst Appl 92:507–520
16. Tarus JK, Niu Z, Yousif A (2017) A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. Future Gen Comput Syst 72:37–48
17. Frémal S, Lecron F (2017) Weighting strategies for a recommender system using item clustering based on genres. Exp Syst Appl 77:105–113
18. Zhang L, Yang S, Zhang M (2014) E-commerce website recommender system based on dissimilarity and association rule. TELKOMNIKA Indonesian J Electric Eng 12(1):353–360
19. Patro SGK, Mishra BK, Panda SK, Kumar R, Long HV, Taniar D, Priyadarshini I (2020) A hybrid action-related K-nearest neighbour (HAR-KNN) approach for recommendation systems. IEEE Access 8:90978–90991
20. Lv J, Song B, Guo J, Du X, Guizani M (2019) Interest-related item similarity model based on multimodal data for top-N recommendation. IEEE Access
21. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y (2018) DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Method 18:1–12

Improved Ant Colony Optimization for Optimal Resource Utilization in Cloud Computing

Diwakar Bhardwaj, Avadhesh Kumar Gupta, and Ashish Sharma

Abstract Recently, cloud computing is an evolving research field deployed for computing by many researchers. The computing is offered as a service in cloud which is regarded as novel technology. In order to meet customer necessitates, various services are offered on the basis of customer dynamic request continuously in cloud computing, and it is the foremost task of cloud computing for providing the desired services to every consumer. The challenge lies in servicing all the customers with the limited existing resource, and it has been tricky to meet all the demanded services by the cloud providers. The allotment of perspective cloud resources through the cloud providers is yet another endeavor which should be done in reasonable way. Hence, cloud consumers' quality of service and fulfillment are the most noteworthy factors to be considered. The prevailing research discussed about the challenges, techniques, key performance issues etc., encompassed in cloud computing resource sharing. Ant colony optimization algorithm is greatly utilized for optimizer analysis of load on physical machine on the basis of local migration agent which aids in migrated VMs load computation and for choosing proper physical server. Conversely, trapping of local optima may happen at certain time which in turn impacts on performance degradation pertaining to global search. The search diversity enhancing is one among the possible solutions for evading the trapping into local optima in ACO. Mutation-based improved ant colony optimization (IACO) is greatly deployed in this research work for analysis of physical machine load VM migration besides effectual resource exploitation.

Keywords Physical machine · Ant colony optimization · VM migration and resource allocation

D. Bhardwaj (✉) · A. Sharma
Department of Computer Engineering and Application, GLA University, Mathura, India
e-mail: diwakar.bhardwaj@gla.ac.in

A. Sharma
e-mail: ashishs.sharma@gla.ac.in

A. K. Gupta
USCI, Karnavati University, Gujrat, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_38

1 Introduction

In the domain of cloud computing, the sustainment of service-level agreement (SLA) has assured to the customers to facilitate them with resources either as service or as per their requirements or both. Besides, it enables the customers to access their resources according to their requirements. However, there is a possibility of inappropriate resource distribution that leads to squandering of resources, as the cloud computing resources are openly accessible. Whereas, the dynamic resource allocation has employed in the cloud services of infrastructure-as-a-service (IaaS) through the system of parallel data processing [1, 2]. In the processes of cloud, this system shows its adequacy to efficiently parallelize the receiving tasks through large data. Moreover, the Virtualization has primarily intended to empower the solutions based on "Computing-as-a-Service" vision of cloud. The associated features of VM (i.e., flexible resource provisioning and migration) enhance the effectiveness of VM in the dynamic resource provisioning and the resource usage. During the scenarios of overload, the "state" of VM has transmitted by the live VM, from a physical machine to another machine [2–8].

A controller handles a set of host machines regarding to the allocation of VM with energy efficiency deprived of SLA violation. For every VM, the historical data has fragmented to a defined length, which has involved constructing the forecast systems. In general, the migration has processed in two methods, i.e., the temporary suspension of VM before transference and live migration process, where the VM has solely stopped for a split second. The recent study confers the significant issues of performance, the practical difficulties, and the appropriate resource allocation methodologies for cloud computing. Here, the physical machine's load has been assessed by the optimizer with the help of ant colony optimization algorithm, through which a load of VMs has been estimated by the local migration agent for being migrated and choosing a proper physical server. Nonetheless, this process may be trapped by the local optima in such scenarios. Besides, this work presents enhanced ant colony optimization based on mutation, which assesses the load over the physical machine in terms of evading the trap of local optima in ACO, through which the VM has migrated; besides, the resource has been utilized in an efficient way.

Basu et al. [9] tend to diminish the usage of power and the VM scheduling by proposing an enhanced genetic algorithm that enables the nodes to maintain their load, i.e., the loading of nodes will not exceed or decrease their capacity. Here, the consideration for a node can be each chromosome of the population. Subsequently, each VM has been assigned to a node, which is associated with the genes of a chromosome. Post processing of crossover and mutation, the optimization strategies have been applied as regards the accomplishment of further allocation processes. The suggested method shows its proficiency to surpass the previous methods in balancing the load and in the utilization of resource.

Khan et al. [10] contemplated several methodologies for scheduling of VM, besides introduced many approaches for resource allocation. To furnish the hassle-free implementation, and enable the availability of cloud-related datasets in VM, this

study proposes the method that deliberates numerous resource management units. So, this strategy incorporates several workloads; moreover, the usage of servers has been expanded in an effective manner.

Jain et al. [11] intended to investigate a few cloud consolidation and migration approaches, like ant colony optimization, and K-means regression, and furnished the summary on obtained results, for which they proposed an approach that allocates VMs to physical machines (PMs) by employing initial allocation algorithm. Post optimization of physical machines, the migration has processed through ACO, during which the VM has migrated to nearby PM. So, the bandwidth and time have been preserved.

2 Proposed Methodology

In this segment, the proposed framework has been comprehensively conferred, in which the prediction of user's trust value has been performed by improved ACO. In Fig. 1, the typical representation of the proposed framework has been demonstrated.

Following steps are significantly involved in the process of the proposed framework. The structural design of cloud consists of numerous physical hosts with the ability to add maximum VMs whichever is running over it. Hence, the reciprocal correlation of physical host, and VM has sustained during the migration of VM processing through one PM to other PM. On the basis of proposed modules, the VM migration has been processed. Subsequently, the live VM migration has started by the derived output of the proposed module. At the PM of server in the cloud, the proposed modules have executed. In the logical phases of VM migration, its safety and handling of failures have to be ensured, for which some necessary steps have to be done as follows:

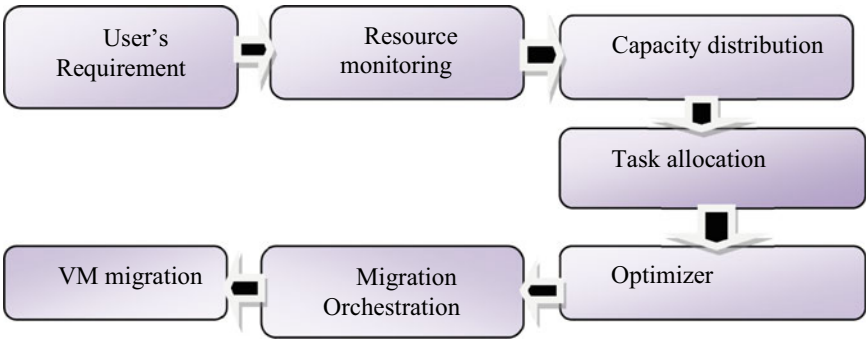


Fig. 1 Overall architecture of the proposed model

Phase 1. Pre-migration: At the time of VM running on the physical machine, this phase begins, during which the destination PM has identified, then it has selected for minimizing the time for overall migration.
Phase 2. Reservation: Triggering of migration request from Physical Machine-A to Physical Machine-B is performed; at that time, availability of necessary resources has ensured to accomplish the VM. If available, the resources will be preserved, or else VM remains to run on Physical Machine-A.
Phase 3. Iterative pre-copy: Here, the overall memory pages have been transmitted from PM-A to PM-B at the first iteration. From the further iterations, the remaining pages have been transferred from previous iterations.
Phase 4. Stop-and-copy: VM from source Physical Machine-A has stopped, and its network traffic has shifted to destination Physical Machine-B. Then, remaining memory pages besides CPU state have been diverted to host B. Though, the VM has been suspended from host A, yet it reflects its primary copy in the scenarios of failure.
Phase 5. Commitment: During this phase, host B acknowledges host A with the successful reception of VM, through which the achievement of migration has been assured. Post reception of successful acknowledgment from host B, the copy of VMs have been removed from host A. Thus, host B turns out to be Primary Machine for VM.
Phase 6. Activation: During this phase, consequent to aforementioned phases, the VM shows its successful run on host B; besides, the VMs directly migrated from overloaded PM to under-loaded PM through the modules of the proposed method.

2.1 Resource Monitoring

The assortment of cloud resources CPU as well as memory usage profile is mainly accomplished through the resource monitoring module. The physical machine's hypervisor communication as well as resource usage collection is greatly achieved by psutillibrary (python system and process utilities) and psutil.cpu_percent is exploited for assessment of CPU deployment of all virtual machine plus physical machine in direct approach. The aggregation of CPU usages of its virtual machines is done for CPU usage profile assessment for physical machine. The cloud resources memory utilization is achieved by process. Memory_info() and saving is done on behalf of future database.

2.2 Capacity Distribution

The distribution of capacity amid the physical machines is attained through capacity distribution which aids in fixing the upper threshold and lower threshold values for the physical machines, and it also supports for load detection in optimizer module.

2.3 Task Allocation

The allotment of machine task is accomplished by means of computation of execution time besides memory required for the next task which is the aim of task allocation. This might greatly support for migration phase on the basis of task calculations distribution and thereby diminishing the migration time.

2.4 Optimizer

Resource monitor is primarily deployed for receiving the resource usage profiles through the optimizer. The physical machine load amounts which are in active state is mainly determined by accomplishing ACO algorithm. A fresh list of virtual machines allocation for physical machines is obtained by effective execution of enhanced ACO-based VM distribution list algorithm. With local migration agent support, ensuing VM redistribution message on PM is communicated to the migration orchestration.

2.5 Optimizer Analyses Load on Physical Machine Using Improved Ant Colony Optimization

2.5.1 Graph Representation

After completing the reformulation of feature selection issue into an ACO-suitable problem, the ACO necessitates the problematic representation in terms of graph. At this time nodes denote features, with the edges amid them representing the next feature choice. An ant traversal through the graph where a lowest number of nodes are visited satisfying traversal stopping criterion leads for searching the optimal feature subset. There is no full connection of nodes for permitting a few features to be designated later. The ant is presently at node $f1$ and besides has an option of which feature to be added next to its path (dotted lines). The feature $f2$ is chosen next on the basis of transition rule, then $f3$ plus then $f4$. Upon arrival at $f4$, the assessment of current subset $\{f1, f2, f3, f4\}$ is done for substantial the traversal stopping criterion (e.g., appropriately enhanced classification accuracy has been attained expending

this subset). The ant ends its traversal and yields this feature subset as a candidate for data reduction.

The graph representation reformulation forms the basis for transition rules and pheromone update standard ACO algorithms rules which might be applied. Here, there is no link amid pheromone and heuristic value, and instead, every feature possess its specific pheromone value besides heuristic value [12–14].

2.5.2 Heuristic Desirability

ACO algorithm is generally regarded as constructive heuristic for probabilistically constructing elucidations. An element sequences obtained from the solution components which is of finite set are assembled by means of constructive heuristic. The commencement of solution construction is done with an empty partial solution. A feasible solution component inclusion is performed through extension of present partial solution at every construction step. A precise heuristic interest of traversing amid features may perhaps be a few subset assessment function, for instance, an entropy-based measure otherwise rough set dependency measure. The heuristic information for feature selection is chiefly referred to the classifier performance in this algorithm. The amalgamation of heuristic desirability of traversal and node pheromone levels is termed as probabilistic transition rule, signifying the probability that ant k will take account of feature i in its solution at time step t:

$$P_i^k(t) = \begin{cases} \frac{[\tau_i(t)]^x.[\eta_i]^\beta}{\sum_{u \in j\,k}[\tau_u(t)]^x.[\eta_u]^\beta} & \text{if } i \in j^k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where J^k represents set of feasible features that might be included to the partial solution; si as well as gi are correspondingly the pheromone value and heuristic desirability accompanying with feature i. a and b are two factors signifying the relative prominence of the pheromone value and heuristic information.

(A) Real ants follow a path amid nest besides food source.
(B) An obstacle appears on the path: Ants choose whether to turn left or right with identical probability.
(C) Rapid deposition of pheromone is done on the shorter path.
(D) The shorter path is selected by every ant.

The balance amid pheromone intensity (i.e., previous successful moves history), si, besides heuristic information (expressing desirability of the move), gi is attained by transition probability used by ACO. Thereby, achieving the exploitation–exploration trade-off in an effectual way. Utilizing the previous search process favors actions which are regarded as an effective one for manipulating information attained about the search space. The examination of previously unseen actions during the search for identifying such actions is also performed for search space exploration. The right assortment of the parameters a and b is chiefly utilized for improved balancing amid

exploitation and exploration. The previous search experience is made negligible whenever is $a = 0$, referring no pheromone information is exploited and degradation of search to stochastic greedy search happens. When $b = 0$, the attractiveness (or potential benefit) of moves is made negligible.

2.5.3 Pheromone Update Rule

The pheromone evaporation on all nodes triggering is accomplished once entire ant solutions are concluded and based on Eq. (3),pheromone depositing by each ant k is performed, $\Delta\tau_i^k(t)$, on all the used nodes

$$\Delta\tau_i^k(t) = \begin{cases} \emptyset.\gamma\left(s^k(t)\right) + \frac{\varphi.\left(n-\left|s^k(t)\right|\right)}{n} & \text{if } i \in s^k(t) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $S^k(t)$ denotes feature subset determined by ant k at iteration t, and $\left|s^k(t)\right|$ represents its length. On the basis of classifier performance measure, $\gamma\left(s^k(t)\right)$ along with feature subset length $\emptyset$ updating of pheromone is done. Also, the relative weight of classifier performance and feature subset length $\emptyset \in [0, 1]$ and $\varphi = 1 - \emptyset$ are controlled by utilizing the parameters $\emptyset$ and u φ. The feature selection task tends to exhibit dissimilar impact on classifier performance, and besides, feature subset is inferred through formula. It is presumed that classifier performance is more noteworthy than subset length in this experimentation, hence fix/ $= 0.8$, $u = 0.2$. The rule for pheromone evaporation implementation and new pheromone inclusion by ants is practically given by

$$\tau_i(t+1) = (1-\rho)\tau_i(t) + \sum_{k=1}^{m} \Delta\tau_i^k(t) + \Delta\tau_i^g(t) \tag{3}$$

where m represents number of ants at every iteration and $q2$ (0,1) notates the pheromone trail decay coefficient. Stagnation is main objective of performing pheromone evaporation stating that all ants building the same solution. g indicates the best ant at every iteration, during which updating of pheromone based on all ants is given in Eq. 4 as well as the deposition of best ant for further pheromone is done on best solution nodes. The exploration of ants around the optimum solution in subsequent iterations is achieved thereby.

2.5.4 Solution Construction

ACO feature selection along with the complete process is revealed in Fig. 2 which starts through generation of number of ants as well as positioning arbitrarily on the graph which possess distinct random feature. On the other hand, there exists

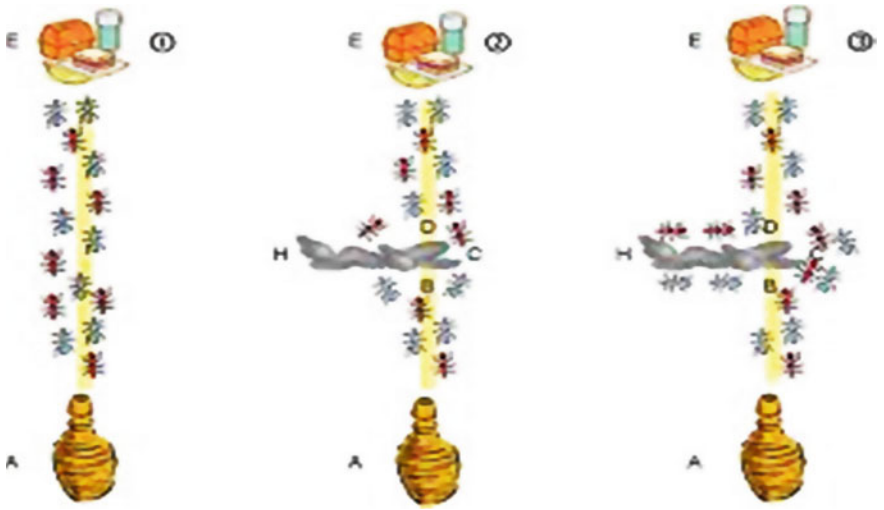


Fig. 2 ACO working procedure

identical number of ants on the graph along with the number of features contained via data. The path creation is accomplished by the ant utilizing various dissimilar features. They start traversing nodes probabilistic manner from primary position till a traversal stopping criterion is met. The assessment is done on the basis of congregated subsets. The process termination takes place whenever an optimal subset exist or else execution of algorithm happens definite number of times and besides yields the best feature subset. The updating of pheromone is done and if not even a single conditions is met, the iteration of process reinitiates after the creation of new set of ants.

2.5.5 Improved ACO

Ant colony optimization algorithm with mutation by means of self-adaptive approach is greatly utilized in this research work. The notion of mutation is utilized for algorithm improvement through local optima escaping. ACO is mainly involved for all VM visiting through making a tour $T\ (i, j)$. Through utilizing new mutation rate created via self-adaptive methodology, generation of new tour $T_{\text{new}}\ (i, j)$ is done through application of supplementary mutation operator on the tour $T\ (i, j)$. The comparison of new tour $T_{\text{new}}\ (i, j)$ with $T\ (i, j)$ is accomplished thereby and followed by substitution of $T\ (i, j)$ using $T_{\text{new}}\ (i, j)$ takes place when cost of new tour is not as much than the cost of $T\ (i, j)$. The repetition of this procedure is done till maximum iteration is not attained (Fig. 3).

The key steps of the commended algorithm are stated below:

1. Activation.
   - Obtain ant population.

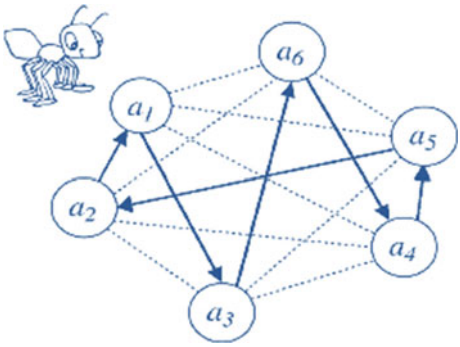


Fig. 3 Mutation-based ant colony optimization

- Define the pheromone trial intensity allied to any feature.
- Estimate the highest number of permitted iterations

2. Generation of solution and assessment of ants.
   - Designate any ant arbitrarily to a single feature alongside visiting features, through which each ant enables to construct over all solutions. During this phase, the classifier's mean square error (MSE) has known to be the assessment factor. If an ant is incapable of decreasing the classifier MSE within ten consecutive attempts, it will depart once after completing its task.
3. Assessment of the selected subsets.
   - Categorize selected subsets on the basis of classifier performance as well as their length. At that point, choose the optimal subset.
4. Verify the stop criterion.
   - Exit, if the iteration count exceeds the highest number of permitted iterations, else resume the steps.
5. Carry out local trail update;
   If mutation criteria has encountered, then choose random VM from present tour $T\ (i, t)$ for mutation operator. Involve mutation operation for obtaining new VM new tour $T_{\text{new}}\ (i, t)$ created from the mutation outcome.
6. Generation of new ants.
   - Proceed to remove the preceding ants and generate new ants.
7. Proceed and continue from step 2.

2.5.6 Local Migration Agent

All VMs are rearranged through the local migration agent from PM instance on the basis of their moderate load, as soon as an instance load condition of the physical server machine exceeds the value of upper threshold.

As expressed by the following equation, this load assessment has accomplished, in which the higher loaded VM has significantly prioritized.

$$\overline{\mathrm{Vi}(j,T)} = \frac{1}{T}\sum_{k=1}^{n} \mathrm{vi}(j,k)(\mathrm{tk} - \mathrm{tk} - 1) \tag{4}$$

In the time period K, the jth instance of VMon the ith instance of PM obtains the resource, which has denoted by Vi (j, k).

The VM migration has organized by the migration orchestration, which has employed to directly process the actual VMs migration.

3 Result and Discussions

The evaluation of the recommended method is accomplished on the basis of modeling random workload through experimentation by utilizing four physical machines whose specification are 8 GB RAM and Core i5-6200U. These physical machines are utilized for creation of three ESX_i virtual machines whose specification are 4 GB RAM and amid which one physical machine serves as V_{center} Server appliance, other for NAS storage installation. The generation of CPU consumption is done stochastically for demonstration of proposed technique preeminence. The various performance metrics such as time, throughput and resource utilization are contrasted for the proposed IACO with that of the prevailing IGA and ACO approaches (Table 1).

The various classification approaches with their throughput is contrasted for the suggested IACO with that of the prevailing IGA and ACO approaches for validating the performance which is shown in Fig. 4. It is inferred that superior throughput results of 550 (kbps) are obtained for suggested IACO, while the prevailing IGA and ACO approaches give only 400 (kbps) and 500 (kbps), respectively.

The various classification approaches with their migration time is contrasted for the suggested IACO with that of the prevailing IGA and ACO approaches for validating the performance which is shown in Fig. 5. It is inferred that lower time of 320 (sec) is utilized for suggested IACO while the prevailing IGA and ACO approaches takes 600 (sec) and 500 (sec), respectively.

The various classification approaches with their resource utilization is contrasted for the suggested IACO with that of the prevailing IGA and ACO approaches for validating the performance which is shown in Fig. 6. It is inferred that greater resource

Table 1 Performance comparison results

Metrics	Methods IGA	ACO	IACO
Throughput (Kbps)	400 (Kbps)	500 (Kbps)	550 (Kbps)
Migration time (sec)	600 (sec)	500 (sec)	320 (sec)
Resource utilization	0.7	0.8	0.9

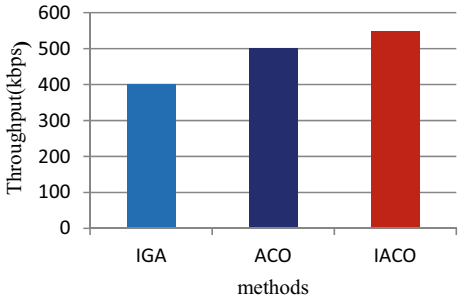


Fig. 4 Throughput results verses classification methods

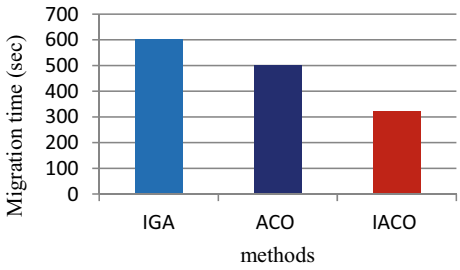


Fig. 5 Migration time results verses classification methods

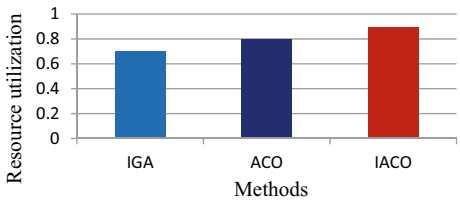


Fig. 6 Resource utilization results verses classification methods

utilization of 0.9 is obtained, while the prevailing IGA and ACO approaches yield 0.7 and 0.8, respectively.

4 Conclusion and Future Work

Cloud computing is one among the evolving research fields deployed for computing. In order to meet customer necessitates, various services are offered on the basis of customer dynamic request continuously in cloud computing, and it is the foremost task of cloud computing for providing the desired services to every consumer. The challenge lies in servicing all the customers with the limited existing resource, and it has been tricky to meet all the demanded services by the cloud providers. The allotment of perspective cloud resources through the cloud providers is yet another endeavor which should be done in reasonable way. In this work, improved

ant colony optimization is exploited for mitigating the resource allotment issues in cloud computing environment. Nonetheless, a security issue in cloud is another eminent challenge and ought to be concentrated in future work through utilization of some frame work.

References

1. Vanitha M, Marikkannu P (2017) Effective resource utilization in cloud environment through a dynamic well-organized load balancing algorithm for virtual machines. Comput Electr Eng 57:199–208
2. Shah NB, Shah ND, Bhatia J, Trivedi H (2019) Profiling-based effective resource utilization in cloud environment using divide and conquer method. Information and communication technology for competitive strategies. Springer, Singapore, pp 495–508
3. Kushwah VS, Goyal SK, Sharma A (2020) Maximize resource utilization using aco in cloud computing environment for load balancing. Soft computing: theories and applications. Springer, Singapore, pp 583–590
4. Gupta J, Azharuddin M, Jana PK (2016) An effective task scheduling approach for cloud computing environment. In: Proceedings of the international conference on signal, networks, computing, and systems. Springer, New Delhi, pp 163–169
5. Chaturvedi A, Kumar R (2021) Multipath routing using improved grey wolf optimizer (IGWO)-Based Ad Hoc on-Demand Distance Vector Routing (AODV) Algorithm on MANET. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
6. Kumar R, Chaturvedi A (2021) Improved cuckoo search with artificial bee colony for efficient load balancing in cloud computing environment. In: Smart innovations in communication and computational sciences. Advances in intelligent systems and computing, vol 1168
7. Kumar R, Bhardwaj D, Mishra MK (2020) Enhance the lifespan of underwater sensor network through energy efficient hybrid data communication scheme. In: 2020 international conference on power electronics & IoT applications in renewable energy and its control (PARC). Mathura, Uttar Pradesh, India, pp 355–359
8. Shukla DK, Dwivedi VK, Trivedi MC (2020) Encryption algorithms in cloud computing. In: Elsevier's journal -materials today proceedings
9. Basu S, Kannayaram G, Ramasubbareddy S, Venkatasubbaiah C (2019) Improved genetic algorithm for monitoring of virtual machines in cloud environment. Smart intelligent computing and applications. Springer, Singapore, pp 319–326
10. Khan H, Janjua K, Sikandar A, Qazi MW, Hameed Z (2020) An Efficient Scheduling based cloud computing technique using virtual Machine Resource Allocation for efficient resource utilization of Servers. In: 2020 international conference on engineering and emerging technologies (ICEET). IEEE, pp 1–7
11. Jain S, Dhoot K, Rede A, Adeshara N, Mhamane S (2019) Optimization of resources in cloud computing using virtual machine consolidation. In: 2019 international conference on smart systems and inventive technology (ICSSIT). IEEE, pp 1285–1288
12. Yang Q, Chen WN, Yu Z, Gu T, Li Y, Zhang H, Zhang J (2016) Adaptive multimodal continuous ant colony optimization. IEEE Trans Evol Comput 21(2):191–205
13. Kulkarni PA (2019) Explore-exploit-explore in ant colony optimization. In: Proceedings of the 2nd international conference on data engineering and communication technology. Springer, Singapore, pp 183–189
14. Raviprabakaran V, Subramanian RC (2018) Enhanced ant colony optimization to solve the optimal power flow with ecological emission. Int J Syst Assur Eng Manage 9(1):58–65

An Improved *K*-Means Clustering with Convolutional Neural Network for Financial Crisis Prediction

Aruna Dhamija, Deepika Pandoi, Krishanveer Singh, and Sachin Malhotra

Abstract A great interest is shown by financial firms and institutions in predicting organization's financial failures in recent days. Taking proper decisions by financial organizations at right time are assisted using an important area called financial crisis prediction (FCP). In companies, wrong decision-making may lead to bankruptcy or financial crisis and clients, vendors, and investors may be affected. In statistical techniques, various works applied for solving FCP problem to review earlier techniques. However, highly intelligent and automated prediction alone is not enough. For better classification performance, an effective prediction model should be developed and if should be able adapt various datasets. So, for predicting financial crisis, an improved *K*-means clustering and convolutional neural network (CNN) are designed in proposed system. There are two stages in designed system, namely classification stage based on convolutional neural network (CNN) and clustering based on improved *K*-means. Wrongly clustered data are eliminated using a proposed an improved *K*-means algorithm. In second stage, financial data instances are classified using convolutional neural network (CNN) for identifying whether the financial firm will undergo failure or not. Three benchmark datasets, namely Polish, Weislaw, and bankruptcy, are used in experimentation for validating proposed algorithm's performance. With respect to F-measure, false positive rate (FPR), and accuracy, better performance is attained using proposed technique as shown in experimental results.

Keywords False positive rate (FPR) · Convolutional neural network (CNN) · Improved *K*-means clustering · Financial crisis prediction (FCP)

A. Dhamija (✉) · D. Pandoi · K. Singh
Institute of Business Management, GLA University, Mathura 281406, India
e-mail: aruna.dhamija@gla.ac.in

D. Pandoi
e-mail: deepika.pandoi@gla.ac.in

K. Singh
e-mail: krishanveer.singh@gla.ac.in

S. Malhotra
Chandigarh University, Chandigarh, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_39

1 Introduction

Financial crisis prediction (FCP) has attracted much attention in the past decade due to raise in companies financial crisis in every part of the world [1]. For financial firms, it is highly essential to have financial crisis prediction (FCP), where possible risk is computed for reducing upcoming losses and if a default risk is greater than predefined acceptance level, then new credit proposals should be avoided [2, 3].

This process is also termed as credit default classification process, in which customer is referred as "non-default," if he pays back the loan, otherwise he is referred as "default." In computing profitability and productivity of financial firm, an important role is played by FCP's accuracy [4]. For example, organization's huge future loss can be minimized by a small accuracy level adjustment of potential user using default credit.

Using intelligent models or statistical techniques or mathematics theory, based on current financial condition, financial crisis of a firm is predicted using FCP. A type of binary classification is performed in this FCP, where crisis and healthy firms are distinguished. For a model, in real-world life, model validity is reflected using an important factor called prediction accuracy. Models having good prediction accuracy will possess practical application value. So, model's prediction accuracy enhancement is a major problem of FCP.

Bankruptcy and credit-related properties debt rating's degradation causes financial crisis. In past years, default techniques are used, but financial crisis in 2007/2008 leads to requirement of effective FCP models with utmost priority. For corporate FCP, there is no standard models or theories as suggested in [5]. The theoretical model absence for investigating financial crisis to exploratory actions for discriminant features identification and prediction models using trial and error [6, 7].

Using intelligent models or statistical techniques or mathematics theory, based on current financial condition, financial crisis of a firm is predicted using FCP. A type of binary classification is performed in this FCP, where crisis and healthy firms are distinguished. For a model, in real-world life, model validity is reflected using an important factor called prediction accuracy. Models having good prediction accuracy will possess practical application value. So, model's prediction accuracy enhancement is a major problem of FCP. For a model, in real-world life, model validity is reflected using an important factor called prediction accuracy. Models having good prediction accuracy will possess practical application value. So, model's prediction accuracy enhancement is a major problem of FCP. A type of binary classification is performed in this FCP, where crisis and healthy firms are distinguished. Models having better prediction accuracy will only possess practical application. So, the major concern is that how to enhance FCP models prediction accuracy whether a firm will face financial crisis or not according to current financial condition using mathematics theory. Various quantitative models are used by professional and academicians for enhancing FCP performance. Two novel kernels are developed by Reddy and Ravi [8] using soft computing techniques for data classification [8]. For FCP, efficient results are produced by presented technique as shown in experimental results.

Nelder–Mead simplex and ant colony optimization (ACO) are used by Sharma et al. [9] to present a hybrid technique for training NN with an FPC application [9]. A new combination model is presented by Wang and Wu [10], where PNN is combined with back-propagation (BP) based on adaptive boosting algorithm for financial crisis prediction. Financial crisis can be predicted in an excellent manner using this new combination model [10]. Sample data collected from 100 listed Taiwan companies are used by Chen [11] for performing corporate FCP. Along various classifiers like logistic regression (LR), decision tree (DT), appropriate variables are filtered using principal component analysis (PCA) [11]. For FCP, introduced kernel principal component neural network (KPCNN) by Ravisankar and Ravi [12], where significant results are obtained using KPCNN in bank reports [12].

Heo and Yang [13] stated that highly excellent accuracy can be obtained using AdaBoost technique in predicting Korean construction companies financial risk [13]. For performing multivariate diagnosis and forecasting Taiwan companies financial risk, a Mahalanobis-Taguchi system (MTS) is implemented by Lee and Teng [14]. For identifying better FCP mode, a comparison is made with neural network (NN), logistic regression (LR) results [14].

2 Related Work

A classification model based on clustering is introduced in proposed FCP model. It has two stages, namely clustering based on improved *K*-means and classification based on convolutional neural network (CNN). At first, Weislaw, Polish, and qualitative bankruptcy datasets are given as an input. Wrongly clustered data are eliminated using a proposed an improved *K*-means algorithm. Then, designed a model to fit with specified dataset by selecting a rule-based model. At last, rule-based model's optimum parameters are computed using a CNN.

2.1 Improved K-Means Clustering

Observations are partitioned into various clusters using clustering techniques. So, there will be high correlation among observations in same cluster when compared with observation is different clusters. A based clustering technique based on distance called *K*-means is used for clustering, where similarity is computed using distance. High similarity values are shown by objects that are close to each other. Process of *K*-means clustering is described as

1. At first, input is derived from dataset. From available N value, K is selected and it indicates primary cluster centers. Here, $K = 2$.
2. Distance between every object and cluster centers is computed. According to expression (1), objects are joined with nearby cluster.

$$S_i^{(t)} = \{\forall j A_j A_k x_p : \left|\left| x_p - m_i^{(t)} \right|\right|^2 \leq x_p - m_j^{(t)2} \forall 1 \leq j \leq k\} \forall j, A_j A_k \quad (1)$$

3. For ensuring change in cluster center, every cluster center is recomputed as

$$m_i^{(t+1)} = \frac{1}{S_i^{(t)}} \sum_{X_j \in S_i^{(t)}} X_j \quad (2)$$

4. Above two steps are repeated till getting a new cluster center identical to original one.

As there are 2 values in class variables, the K value is selected as 2 in this work. Classification is done using a pre-processed data. In K-means technique, a major problem is in WEKA tool, where cluster performance is straightly influenced by seed values. Some steps are included for eliminating changes due to seed value's random nature in simulation. At first, a program for storing and arranging value known as "within cluster sum of squared errors" in an ascending order.

In every experimentation, "within cluster sum of squared errors" represents a seed value. Effective results are obtained with small value of this. In order to represent 1 to 10,000 seed values, we have stored 10,000 values. In second step, employed the good quality seed values. So, value of 100 is assigned as initial value of seed. Then, mis-clustered data are removed by inserting a loop and rate is determined. It will be shifted to subsequent level, if rate exceeds 75%.

Else, it tries to enter into other seed value by existing from this group. If there is no proper seed to produce a rate value more than 75% after 10,000 or 60,000 iterations, employ, a most proximate rate and corresponding seed are moved to consequent level. Figure 1 shows the improved K-means clustering's flow diagram.

$$\text{Rate} = \frac{\text{Remaining data}}{\text{Sum}} \quad (3)$$

2.2 *Convolutional Neural Network (CNN) for Financial Crisis Prediction*

Financial data instances are classified using convolutional neural network (CNN) model in this proposed research work for identifying whether financial firm by affected by failure or not. Take a classification problem with c classes in n dimensional pattern space and with p vectors $X_i = [x_{i1}, x_{i2},.,x_{in},]$.

Convolutional neural network (CNN) is a highly powerful deep networks, which include multiple hidden layers. From input data, low to high level features are extracted by performing subsampling and convolution operation. There are three

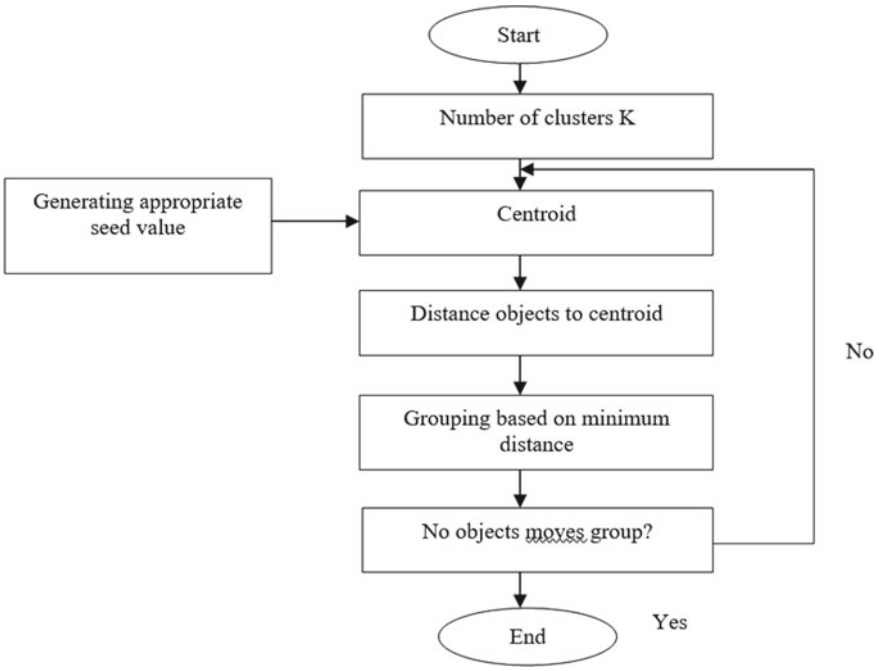


Fig. 1 Flow diagram of improved *K*-means clustering

kind of layers in this networks in general, namely output, input, and hidden layer. Trained output is produced at output layer, and hidden layer is intermediate layer. Figure 2 illustrates the architecture of this CNN.

2.2.1 Convolution Layer

In this layer, input corresponds to clustered attributes and with a kernel (filter), it is convolved. With kernel, input matrix's every block is convolved independently and in output attributes are generated. Output attribute maps are generated using kernel and input attribute's convolution results. In general, filter corresponds to convolution matrix kernel and attribute maps with $i * i$ size corresponds to output attribute computing using convolution of input and kernel.

Multiple convolutional layers are included in CNN, and attributes are produced as next convolutional layers input and outputs. Every convolution layer has bunch of n filters. Input is convolved with these filters and in convolution operation, applied filters count will equalize the generated attribute maps ($n *$) depth.

The $C_j^{(l)}$ represents lth convolution layer's output and it has attributes maps. It is given by

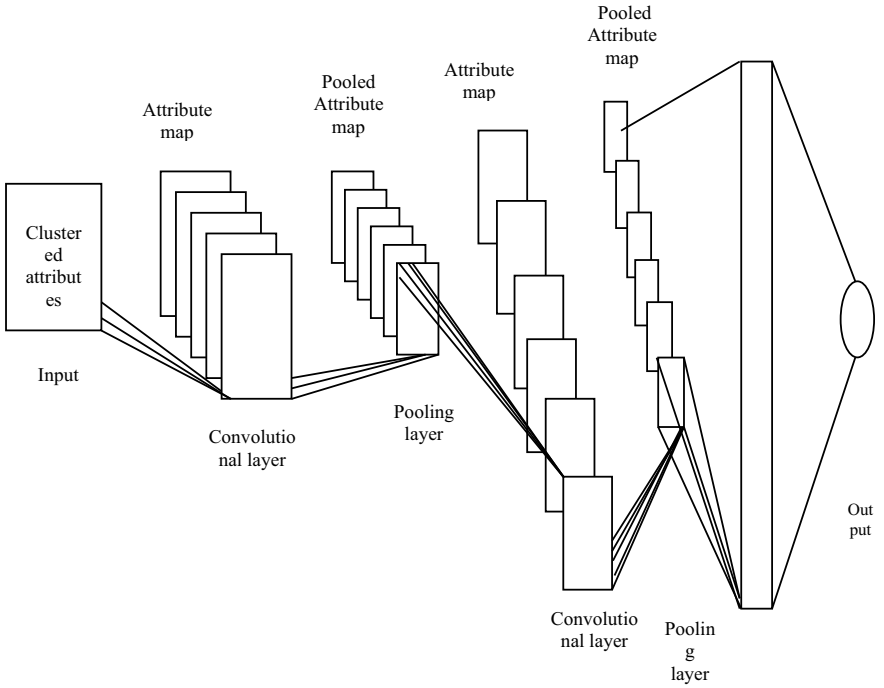


Fig. 2 Convolutional neural network

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l)} \tag{4}$$

where bias matrix is represented as $B_i^{(l)}$ and kernel or convolution filter is represented as $K_{i,j}^{(l-1)}$ and it has $a * a$ size and jth feature map in layer $(l-1)$ and ith attribute map in same layer is connected using this. Attribute maps will be there in output $C_i^{(l)}$ layer. In expression (5), first convolutional layer $C_i^{(l-1)}$ is input space, that is, $C_i^{(0)} = X_i$.

Attribute maps are generated by kernels. After convolution layer, for convolution layer's output's nonlinear transformation, applied the activation function.

$$Y_i^{(l)} = Y\left(C_i^{(l)}\right) \tag{5}$$

where activation function output is represented as $Y_i^{(l)}$ and input is represented as $C_i^{(l)}$.

Rectified linear units (ReLUs), tanh, and sigmoid are the commonly used activation functions. This work utilizes ReLUs activation function, which is represented as $Y_i^{(l)} = \max\left(0, Y_i^{(l)}\right)$. Because of its minimized nonlinear effects and interaction,

this function has wide use in deep learning models. For a positive input value, same value is returned by ReLU and for a negative input value, it produces 0 as output.

Faster training is a major advantage of this activation function due to error derivative, when compared with other functions. In saturating region, it has very small error derivative value and there will be a vanishing of weights. This is termed as vanishing gradient problem.

2.2.2 Subsampling Layer

After convolutional layer, subsampling layer is formed. The attribute maps which are extracted from the previous convolution layer's dimensionality are reduced spatially using this subsampling layer. Selected a mask with $b * b$ size for the same and between attributes map and mask, performed the subsampling operation.

$$X_j^l = f\left(\beta_j^l \text{down}\left(X_j^{l-1}\right) + b_j^l\right) \tag{6}$$

where subsampling function is represented as down $(\cdot)$. In input dataset, over every distinct *n*-by-*n* attributes, this function will some over. So, *n*-times smaller output along both spatial dimension is produced. An additive bias b and multiplicative bias β are provided by every output.

2.2.3 Full Connection

Softmax activation function is used in output layer.

$$Y_i^{(l)} = f\left(z_i^{(l)}\right), \text{ where } z_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_H y_i^{(l-1)} \tag{7}$$

Where attributes weight values are represented as w_H, complete fully connected layer is used for tuning this values for forming every class representation and a transfer function is represented as f, nonlinearity is represented using this transfer function. Financial data's clustered attributes instances are classified using proposed system for identifying, whether financial firm will be effected by failure or not.

3 Experimental Results

With respect to *F*-measure, accuracy, and false positive rate (FPR), comparison between existing improved *K*-means with FSCGACA, fitness-scaling chaotic genetic ant colony algorithm (FSCGACA), genetic ant colony algorithm (GACA),

and proposed improved *K*-means with CNN is made for showing effectiveness of presented FCP model. Using Weislaw dataset, Polish dataset, and qualitative bankruptcy dataset [15, 16], tested the proposed FCP model, for validating consistent performance and showing its adaptability to various datasets.

Table 1 shows the dataset details. From various financial institutions, own dataset is collected for qualitative analysis and benchmark by UCI repository. From Kim and Han (2003), parameters which are used to collect qualitative dataset are derived with attributes like operating risk, competitiveness, credibility, financial flexibility, management risk, and industrial risk. There are 25 instances in qualitative bankruptcy dataset, which includes 107 failed firms and 143 non-failed firms.

There are 43,405 instances in Polish company dataset, which includes 2091 failed firms and 41,314 non-failed firms. Bankrupt Polish companies' details are described in this dataset. During 2000–2012-time period, these companies are investigated and during 2007–2013, validated the operating companies. All5 years data are combined for carry out this study. There are 240 instances in Weislaw dataset, which includes 112 failed firms and 128 non-failed firms.

Removed the incorrectly clustered data by employing improved *K*-means technique and for classifier algorithm, this optimized data are given as an input. Table 2 specified the optimized dataset details. Figure 3 shows the output of improved *K*-means clustering algorithm for applied dataset. Table 3 represents performance comparison.

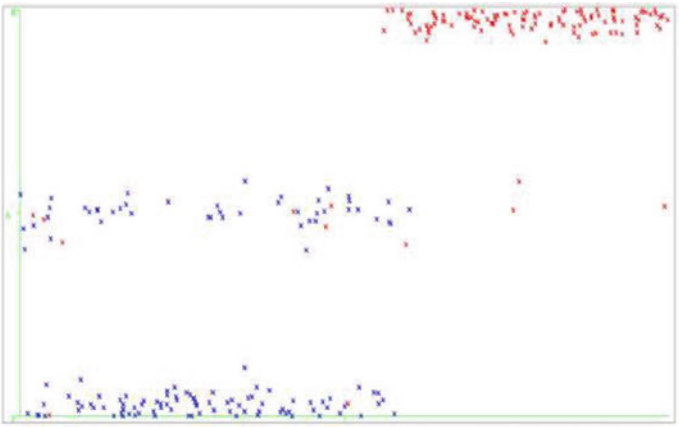
FPR metric comparison between existing improved *K*-means with FSCGACA, FSCGACA, and GACA, proposed improved *K*-means with CNN is shown in Fig. 4. Datasets are represented in x-axis and in y-axis, FPR values are represented. Lower FPR values are produced by proposed *K*-means with CNN when compared with existing techniques for all datasets as indicated in experimental results.

Table 1 Dataset description

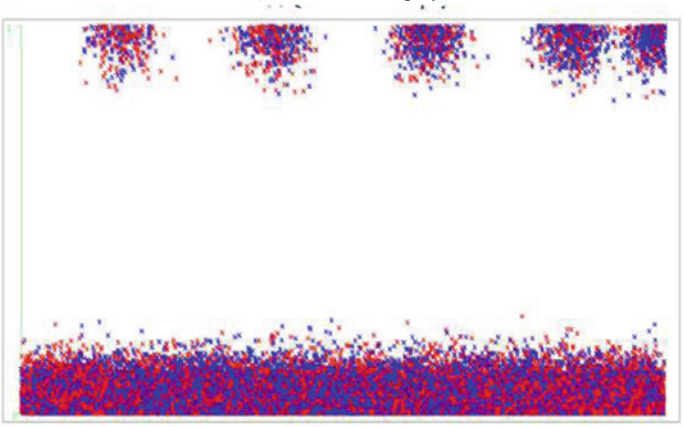
Dataset	Source	# of instance	# of attribute	# of classes	Failed/non-failed
Qualitative	UCI	250	6	2	107/143
Polish companies	UCI	43,405	64	2	2091/41,314
Weislaw	Pietruszkiewicz	240	30	2	112/128

Table 2 Result of *K*-means cluster of initial dataset

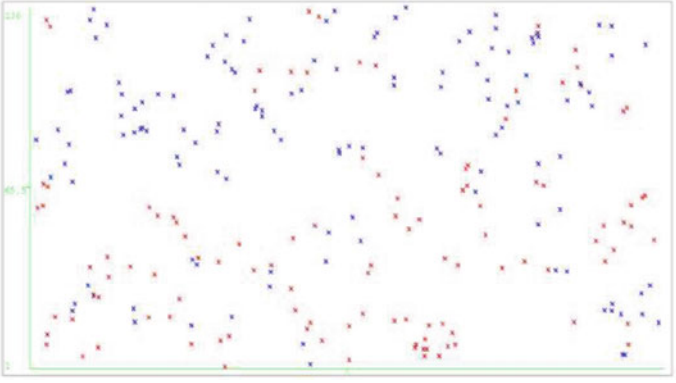
Dataset	Source	# of instance	# of attribute	# of classes	Failed/non failed
Qualitative	UCI	242	6	2	107/135
Polish companies	UCI	36,219	64	2	1759/34,460
Weislaw	Pietruszkiewicz	204	30	2	108/96



a. Qualitative bankruptcy dataset



b. Polish bankruptcy dataset



c. Weislaw dataset

Fig. 3 Clustered data from improved *K*-means algorithm

Table 3 Performance comparison

Methods	Qualitative dataset			Polish companies dataset			Weislaw dataset		
	FPR	Accuracy	*F*-measure	FPR	Accuracy	*F*-measure	FPR	Accuracy	*F*-measure
GACA	6.81	91.32	90.32	0.52	89.98	46.84	5.81	90.74	91.15
FSCGACA	6.72	92.14	91.16	0.41	91.39	51.12	3.44	92.64	93.3
Improved *K*-means with FSCGACA	0.75	97.93	97.69	0.1	94.72	64.39	1.07	97.55	97.72
Improved *K*-means with CNN	0.4	98	98.8	0.5	95.34	68.8	0.6	98.8	98.6

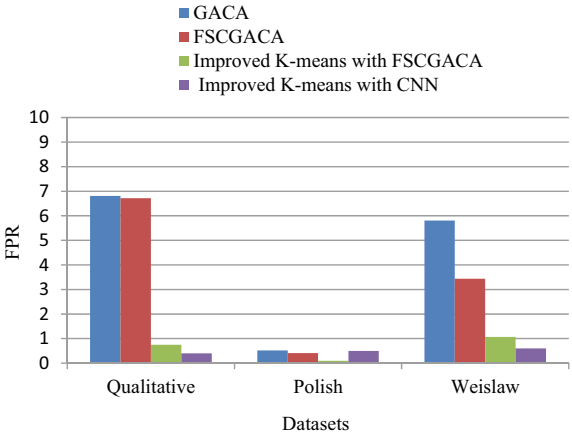


Fig. 4 FPR comparison

Accuracy metric comparison between existing improved *K*-means with FSCGACA, FSCGACA, and GACA, proposed improved *K*-means with CNN is shown in Fig. 5. Datasets are represented in x-axis and in y-axis, accuracy value is represented. High accuracy values are produced by proposed *K*-means with CNN when compared with existing techniques for all datasets as indicated in experimental results.

F-measure metric comparison between existing improved *K*-means with FSCGACA, FSCGACA, and GACA, proposed improved *K*-means with CNN is shown in Fig. 6. Datasets are represented in x-axis and in y-axis, F-measure is represented. High F-measure values are produced by proposed *K*-means with CNN when compared with existing techniques for all datasets as indicated in experimental results.

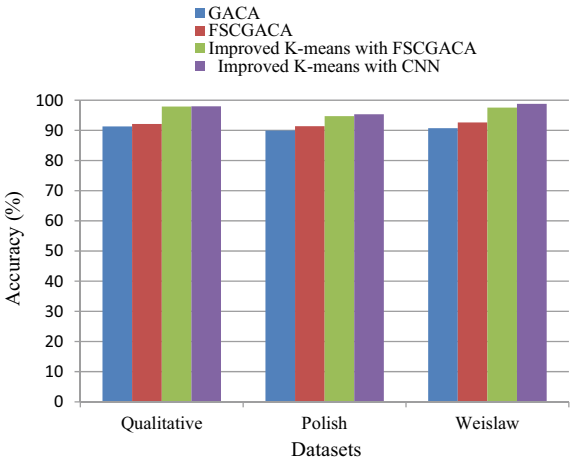


Fig. 5 Accuracy comparison

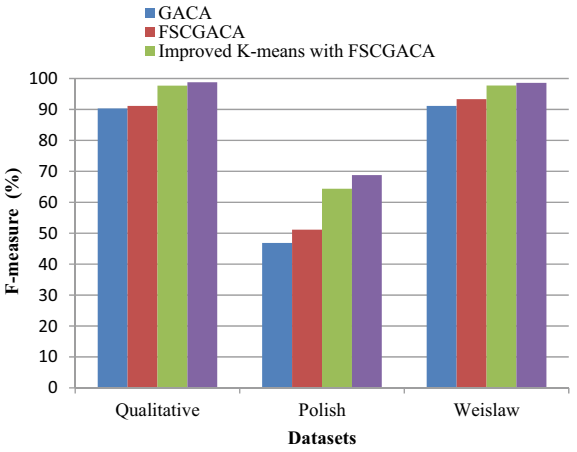


Fig. 6 F-measure comparison

4 Conclusion

For effectively predicting financial crisis, an improved *K*-means clustering with convolutional neural network (CNN) is proposed in this work. At first, datasets are given as input and data points are grouped into clusters by employing improved *K*-means clustering technique. Financial data instance is classified using the convolutional neural network (CNN) for identifying, whether financial firm will face failure or not.

The accuracy of classification is enhanced by integrating improved *K*-means technique with classification. This makes precise classification of financial data. In Weislaw dataset, Polish dataset, and qualitative bankruptcy dataset, high accuracy is produced by proposed FCP model when compared with the previous techniques as shown in experimental results. Ensemble techniques can be used in future for further enhancing proposed FCP model's performance.

References

1. Ala'raj M, Abbod MF (2016) Classifiers consensus system approach for credit scoring. Knowl-Based Syst 104:89–105
2. Wu C, Wang L, Shi J (2015) Financial distress prediction based on support vector machine with a modified kernel function: journal of intelligent systems. J Intell Syst 33(33):177–187
3. Chaudhuri A, De K (2011) Fuzzy support vector machine for bankruptcy prediction. Appl Soft Comput 11(2):2472–2486
4. Acosta-González E, Fernández-Rodríguez F (2014) Forecasting financial failure of firms via genetic algorithms. Comput Econ 43(2):133–157
5. Wang G, Ma J, Yang S (2014) An improved boosting based on feature selection for corporate bankruptcy prediction. Expert Syst Appl 41(5):2353–2361
6. Sun J, Li H (2012) Financial distress prediction using support vector machines: Ensemble vs individual. Appl Soft Comput 12(8):2254–2265

7. Zhou L (2013) Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. Knowl-Based Syst 41:16–25
8. Reddy KN, Ravi V (2013) Differential evolution trained kernel principal component WNN and kernel binary quantile regression: application to banking. Knowl Based Syst 39:45–56
9. Sharma N, Arun N, Ravi V (2013) An ant colony optimisation and Nelder-Mead simplex hybrid algorithm for training neural networks: an application to bankruptcy prediction in banks. Int J Inf Decis Sci 5(2):188–203
10. Wang L, Chong W (2017) A combination of models for financial crisis prediction: integrating probabilistic neural network with back-propagation based on adaptive boosting. Int J Comput Intell Syst 10:507–520
11. Chen MY (2011) Predicting corporate financial distress based on integration of decision tree classification and logistic regression. Expert Syst Appl 38(9):11261–11272
12. Ravisankar P, Ravi V (2009) Failure prediction of banks using threshold accepting trained kernel principal component neural network. In: Proceedings of the IEEE world congress on nature & biologically inspired computing, NaBIC 2009
13. Heo J, Yang JY (2014) AdaBoost based bankruptcy forecasting of Korean construction companies. Appl Soft Comput 24:494–499
14. Lee YC, Teng HL (2009) Predicting the financial crisis by Mahalanobis-Taguchi system–Examples of Taiwan's electronic sector. Expert Syst Appl 36(4):7469–7478
15. Martin A, Uthayakumar J, Nadarajan M (2014) Qualitative_Bankruptcy data set. UCI machine learning repository. https ://archive.ics.uci.edu/ml/datasets/qualitative_bankruptcy. Accessed 1 Oct 2018
16. Pietruszkiewicz W (2008) Dynamical systems and nonlinear Kalman filtering applied in classification. In: Proceedings of the 7th IEEE international conference on cybernetic intelligent systems, CIS 2008

An Empirical Analysis of Multilayer Perceptron Based on Neural Network using Image Segment Dataset

Nidhi Kanwar, Anil Kumar Goswami, and Sachin Malhotra

Abstract Artificial neural network (ANN) is inspired by the biological neurons. Since 1943, it has been unfolded as a favorable system for solving different problems of real-world. ANN consists of many neurons connected in systematic way in input layer, hidden layers, and output layer (Sazli in A brief review of feedforward neural networks. Communications, Faculty of Science, University of Ankara 50(1):1–17, 2006). It learns pattern from the given training data to solve user problems in a similar way as human does. It has the capability of adaptation, generalization, and fault-tolerance (Sharma et al. in International Journal of Advanced research in computer science and software engineering 2(10), 2012). This paper focuses on examining of ANN architectures for data classification and finding optimal model parameters with respect to training and testing accuracy. Thus, the empirical investigation tends to achieve the optimal network architecture and model parameters with minimized training and testing error in multilayer perceptron (MLP) ANN.

Keywords Back propagation algorithm · Artificial neural network (ANN) · Multilayer perceptron (MLP)

1 Introduction

In today's world, with increase in advanced technologies and digitization of the services, the data production (in the form of images, video, and texts) is increasing day by day, which is quite difficult to handle and visualize by a human. There might be some important information in the large amount of data which can be missed from the sight of a human being. There could be some places on earth which human cannot visit

N. Kanwar (✉)
CEA Department, GLA University, Mathura, UP, India

A. K. Goswami
DRDO, Delhi, India

S. Malhotra
Chandigarh University, Chandigarh, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_40

to get information, but this can be done by capturing images of those places and can be further classified. Scientists have devised various data classification approaches and methods whereas every method has different prerequisites and produces output with different level of accuracy [3]. Error can be considered as one of the major constraints for the selection of data classification method, this means selected method must minimize the error to the possible extent. Selection of the classification approach may affect the accuracy of classification; hence, suitable classification method must be selected for the specific study [3]. Due to the raise in highly capable computers and availability of different sources for capturing high quality video sand image sand textsen forced an interest in developing different data classification approaches [4].This study has contributed by empirically examine the ANN architecture for data classification and finding optimal model parameters.

2 Literature Survey

Machine learning (ML) is a subset of artificial intelligence (AI), which is a developed technology that creates, expands, and extends human like intelligence so as to learn, receive, perceive, and process the input data like the human does. "Machine learning is the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions" [5]. According to Arthur Samuel, ML can be defined as a field of study that gives computers the ability to learn without being explicitly programmed. Machines learn with the help of learning algorithms which are fed with some training data to generate the rules based on the knowledge deduction from the given data. It also produces the predictive model for the future use.

Neural network learning is done to make it capable to adapt itself to the changes or stimuli and produce the desired response. Learning is a continuous process which updates its parameters like synaptic weights, bias so that response or output becomes equal to desired output. There are different learning approaches available to train the network like supervised learning, unsupervised learning, reinforcement learning, competitive learning, etc.

"A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output" [6]. In multilayer feed-forward neural network, there should be at least one hidden layer between the other two types of layers, i.e., input layer and the output layer [1]. Different architecture may produce different accuracy for the same data, so the architecture of the network plays a very important role in the accuracy of generalization [7].

In this paper, performance of multilayer perceptron is evaluated using different architectures of the network. There are parameters of the network which can influence its performance like activation function, learning rate, number of hidden layers, number of nodes in each hidden layer, etc. So, this paper shows that how network accuracy increases or decreases with different architectures of the network and it

has been observed that accuracy of the network changes with change in values of parameters.

Many researchers have worked on data classification using different classification approaches. Ertam and Aydin classified the MNIST data using CNN implemented using TensorFlow library [8]. In their experiment, they selected different activation functions in each trial like rectified linear unit (ReLU), exponential linear unit, tanh, sigmoid, softplus, softsign, etc., and found different accuracies using different activation functions. They concluded that highest accuracy, i.e., 98.43% is achieved using ReLU activation function. Campos et al., 2004 implemented the MLP using Object Pascal and Delphi 5.0 as the software implementation tool [9]. The projected network permits the user to give information about the number of processing units in input, hidden, and output layer, the activation function, the output function, the interval required for the random initialization of weights between the layers (the maximum between −1 and 1), the learning rate, the momentum term, and whether bias will be used. The lowest error rate they achieved in their implementation was 3.61 for training dataset and 3.95 for testing dataset which is quite high.

3 Methodology

The proposed method utilized MLP with backpropagation for classification for image segment data. MLP is a supervised algorithm which has targets along with its training data. Initially, the weight parameters of the network initialized randomly to train the network, but these initialized parameters might not give the accurate results as per the user requirements so these parameters need to be updated by some small fraction value to minimize the error.

Accuracy of the network can be calculated by comparing the desired output and the actual output value. This algorithm has a drawback of slow learning which may dominate the learning time of the MLP network.

3.1 Back Propagation Algorithm

The step wise mathematical progress of this algorithm including the updation of weights and calculation of gradient to achieve the optimal network has been illustrated below (Fig. 1).

***Step 1*:** Build a neural network consisting input layer, hidden, and output layer.
***Step 2*:** Provide training file to the network including weight file.
***Step 3*:** Feedforward Pass: Apply this formula to compute the net input of *i*th layer.

$$v_{ji} = (i = 0 \,\text{ton}) w_{ji} x_i \tag{1}$$

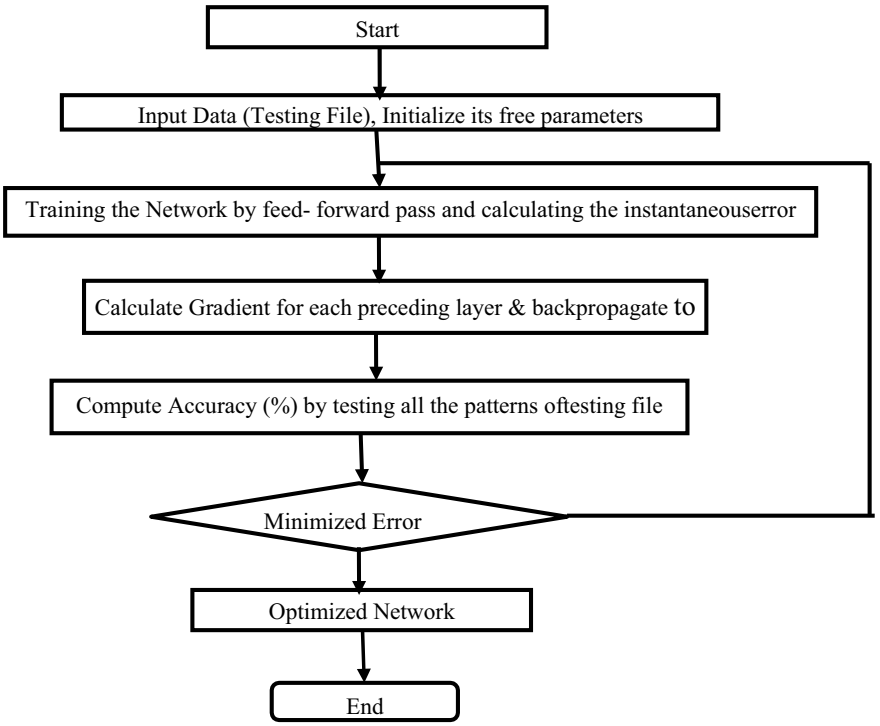


Fig. 1 Flowchart of backpropagation algorithm

***Step 4*:** Apply sigmoid function to the net input to get the output of neuron j

$$y_i = \varphi_j(v_j) \tag{2}$$

***Step 5*:** Calculate squared error function (difference of target output and calculated output)

$$\text{SEF} = \frac{1}{2}\text{Target} - \text{Output} \tag{3}$$

***Step 6*:** If SEF<= Minimum error then go to step 2 else continue with step7.
***Step 7*:** Back-Forward Pass: Compute gradient at output layer L

$$\delta = \eta \quad 1 - \text{output}_j \text{output}_j \text{error}_j \tag{4}$$

where η is learning rate
***Step 8*:** Compute gradient at hidden layers l

$$\delta = \eta 1 - \text{output}_i \text{output}_i \quad w_{jk} \times x_k \tag{5}$$

j= neuron at $i-1$th layer, k= neuron at ith layer

***Step 9*:** Update weights

$$w_{\text{new}} = w_{\text{old}} + \eta \times \delta \times \text{output} \tag{6}$$

***Step 10*:** Repeat steps 3–9

4 Results

MLP is implemented on the following datasets, and their corresponding results have been mentioned below.

4.1 Image Segment Dataset

The dataset with details of feature vectors and their corresponding labels has been described below. The instances were drawn randomly from a database of 7 outdoor images [10]. The images were hand segmented to create a classification for every pixel. Each instance is a 3×3 region. This image segment data have 19 features and 7 class labels. Feature information is shown in Table 1 [10].

Class labels of image segment data are brick face, sky, foliage, cement, window, path, and grass [10]. The training dataset contains 210 samples for training whereas testing dataset contains 2100 samples for the testing of the network. 30 instances per class are taken for training data and 300 instances per class for test data. The file for feature vectors in the form of csv format used for training (Fig. 2).

As we have mentioned above that feature vectors have their corresponding labels which have been shown in Fig. 3.

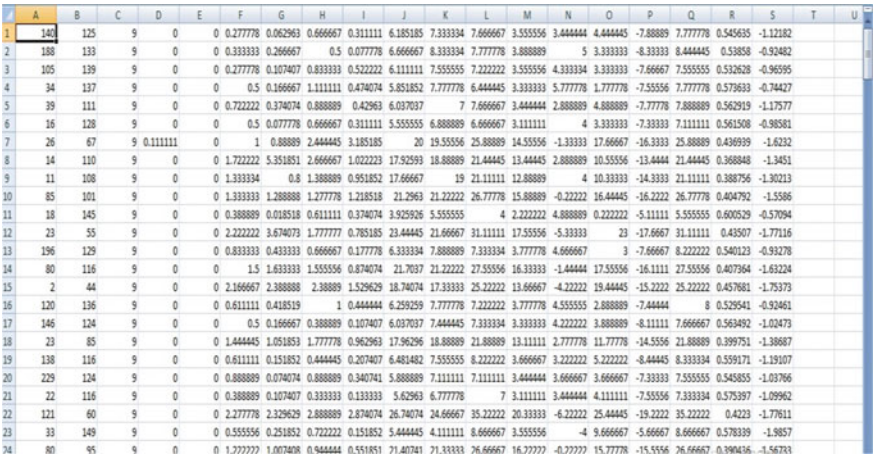
Training: Initially,training option will be chosen by the user to perform feed-forward pass in order to train the network first. After choosing the first option, it will ask user to give values of design parameters such as learning rate, number of hidden layers, and number of nodes in each hidden layer. If feature vectors and their corresponding labels are given in the single csv file then there are two ways to extract features and labels separately. Either user gives the number of inputs and classes or this information is provided in the csv file along with the training data. After giving all this information to the system, it will start training the network and accuracy is also being computed in each iteration.

The training of the network will be continued until the network achieves the minimized error. After the training, updated weight parameters will be displayed on the screen and will be saved in the meta file for testing and generalization.

Results have been shown in the Table 2. The network has been trained by using different architecture and different values for different design parameters. As we have described above about the initialization of design parameters of the network

Table 1 Details of features in dataset

Feature	Feature information
1	Region-ceutroid-col: the column of the center pixel of the region
2	Region-centreid-row: the row of the center pixel of the region
3	Region-pixel-count: the number of pixels in a region = 9
4	Short-line-density-5: the results of a line extraction algorithm that counts how many hues of length 5 (any orientation) with low contrast, less than or equal to 5, go through the region
5	Short-line-density-2: same as short-line-density-5 but counts lines of high contrast, greater than 5
6	Vedge-mean: measure the contrast of horizontally adjacent pixels in the region. There are 6, the mean and standard deviation are given. This attribute is used as a vertical edge detector
7	Vegde-sd: (see 6)
8	Hedge-mean: measures the contrast of vertically adjacent pixels. Used for horizontal hue detection
9	Hedge-sd: (see 8)
10	Intensity-mean: the average over the region of $(R + G + B)/3$
11	Rawred-niean: the average over the region of the R value
12	Rawblue-mean: the average over the region of the B value
13	Rawgreen-mean: the average over the region of the G value
14	Exred-mean: measure the excess red: $(2R - (G + B))$
15	Exblue-mean: measure the excess blue: $(2B - (G + R))$
16	Exgreen-mean: measure the excess green: $(2G - (R + B))$
17	Value-me an: 3D nonlinear transformation of RGB. (Algorithm can be found in Foley and Van Dam, Fundamentals of Interactive Computer Graphics')



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	140	125	9	0	0	0.277778	0.062963	0.666667	0.311111	6.185185	7.333334	7.666667	3.555556	3.444444	4.444445	-7.88889	7.777778	0.545635	-1.12182		
2	188	133	9	0	0	0.333333	0.266667	0.5	0.077778	6.666667	8.333334	7.777778	3.888889	5	3.333333	-8.33333	8.444445	0.53858	-0.92482		
3	105	139	9	0	0	0.277778	0.107407	0.833333	0.522222	6.111111	7.555555	7.222222	3.555556	4.333334	3.333333	-7.66667	7.555555	0.532628	-0.96595		
4	34	137	9	0	0	0.5	0.166667	1.111111	0.474074	5.851852	7.777778	6.444445	3.333333	5.777778	1.777778	-7.55556	7.777778	0.573633	-0.74427		
5	39	111	9	0	0	0.722222	0.374074	0.888889	0.42963	6.037037	7	7.666667	3.444444	2.888889	4.888889	-7.77778	7.888889	0.562919	-1.17577		
6	16	128	9	0	0	0.5	0.077778	0.666667	0.311111	5.555555	6.888889	6.666667	3.111111	4	3.333333	-7.33333	7.111111	0.561508	-0.98581		
7	26	67	9	0.111111	0	1	0.88889	2.444445	3.185185	20	19.55556	25.88889	14.55556	-1.33333	17.66667	-16.3333	25.88889	0.436939	-1.6232		
8	14	110	9	0	0	1.722222	5.351851	2.666667	1.022223	17.92593	18.88889	21.44445	13.44445	2.888889	10.55556	-13.4444	21.44445	0.368848	-1.3451		
9	11	108	9	0	0	1.333334	0.8	1.388889	0.951852	17.66667	19	21.11111	12.88889	4	10.33333	-14.3333	21.11111	0.388756	-1.30213		
10	85	101	9	0	0	1.333333	1.288888	1.277778	1.218518	21.2963	21.22222	26.77778	15.88889	-0.22222	16.44445	-16.2222	26.77778	0.404792	-1.5586		
11	18	145	9	0	0	0.388889	0.018518	0.611111	0.374074	3.925926	5.555555	4	2.222222	4.888889	0.222222	-5.11111	5.555555	0.600529	-0.57094		
12	23	55	9	0	0	2.222222	3.674073	1.777777	0.785185	23.44445	21.66667	31.11111	17.55556	-5.33333	23	-17.6667	31.11111	0.43507	-1.77116		
13	196	129	9	0	0	0.833333	0.433333	0.666667	0.177778	6.333334	7.888889	7.333334	3.777778	4.666667	3	-7.66667	8.222222	0.540123	-0.93278		
14	80	116	9	0	0	1.5	1.633333	1.555556	0.874074	21.7037	21.22222	27.55556	16.33333	-1.44444	17.55556	-16.1111	27.55556	0.407364	-1.63224		
15	2	44	9	0	0	2.166667	2.388888	2.38889	1.529629	18.74074	17.33333	25.22222	13.66667	-4.22222	19.44445	-15.2222	25.22222	0.457681	-1.75373		
16	120	136	9	0	0	0.611111	0.418519	1	0.444444	6.259259	7.777778	7.222222	3.777778	4.555555	2.888889	-7.44444	8	0.529541	-0.92461		
17	146	124	9	0	0	0.5	0.166667	0.388889	0.107407	6.037037	7.444445	7.333334	3.333333	4.222222	3.888889	-8.11111	7.666667	0.563492	-1.02473		
18	23	85	9	0	0	1.444445	1.051853	1.777778	0.962963	17.96296	18.88889	21.88889	13.11111	2.777778	11.77778	-14.5556	21.88889	0.399751	-1.38687		
19	138	116	9	0	0	0.611111	0.151852	0.444445	0.207407	6.481482	7.555555	8.222222	3.666667	3.222222	5.222222	-8.44445	8.333334	0.559171	-1.19107		
20	229	124	9	0	0	0.888889	0.074074	0.888889	0.340741	5.888889	7.111111	7.111111	3.444444	3.666667	3.666667	-7.33333	7.555555	0.545855	-1.03766		
21	22	116	9	0	0	0.388889	0.107407	0.333333	0.133333	5.62963	6.777778	7	3.111111	3.444444	4.111111	-7.55556	7.333334	0.575397	-1.09962		
22	121	60	9	0	0	2.277778	2.329629	2.888889	2.874074	26.74074	24.66667	35.22222	20.33333	-6.22222	25.44445	-19.2222	35.22222	0.4223	-1.77611		
23	33	149	9	0	0	0.555556	0.251852	0.722222	0.151852	5.444445	4.111111	8.666667	3.555556	-4	9.666667	-5.66667	8.666667	0.578339	-1.9857		
24	80	95	9	0	0	1.777777	1.007408	0.944444	0.551851	21.40741	21.33333	26.66667	16.77777	-0.77777	15.77778	-15.5556	26.66667	0.390436	-1.56733		

Fig. 2 Image segment dataset

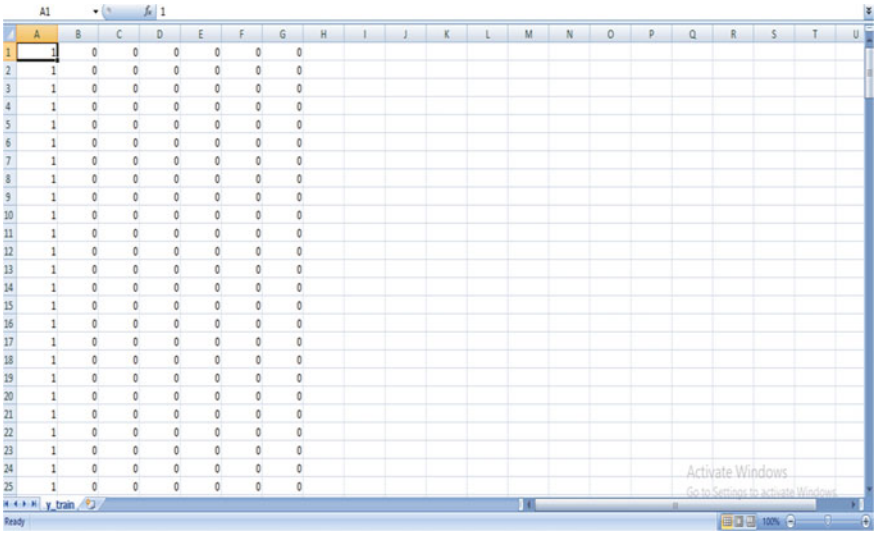


Fig. 3 Labels for the training of MLP

Table 2 Results of image segment data

Learning rate	Number of hidden layers	Number of nodes in each hidden layer	Epochs	Minimized cost	Training accuracy (%)	Testing accuracy (%)
0.1	1	20	2000	0.0025937	80.95	73.59
0.1	1	25	2000	0.0023456	83.34	76.76
0.1	1	30	2000	0.0018143	87.14	81.45
0.1	1	35	2000	0.002401	82.85	76.45
0.1	1	50	2000	0.001156	94.28	87.43
0.1	1	60	2000	0.001652	89.52	82.45
0.1	1	70	2000	0.000799	96.19	89.76
0.1	1	80	2000	0.000908	96.20	87.12
0.1	1	90	2000	0.000899	96.23	87.34
0.1	1	100	2000	0.000697	96.34	89.67
0.1	1	100	1000	0.000676	96.34	89.78

that there are no generalized rules for initializing the design parameters with the right values. So, hit and trial method has been used to find the optimized architecture for the given problem. But according to the results, it can be assumed that increasing the number of hidden neutrons may increase the accuracy of data classification. As it can be seen that highest accuracy had been achieved with 100 neutrons in the hidden layer. And there is no big change between accuracy achieved with 100 and 70 hidden neurons but 100 neurons are giving good accuracy in less iterations.

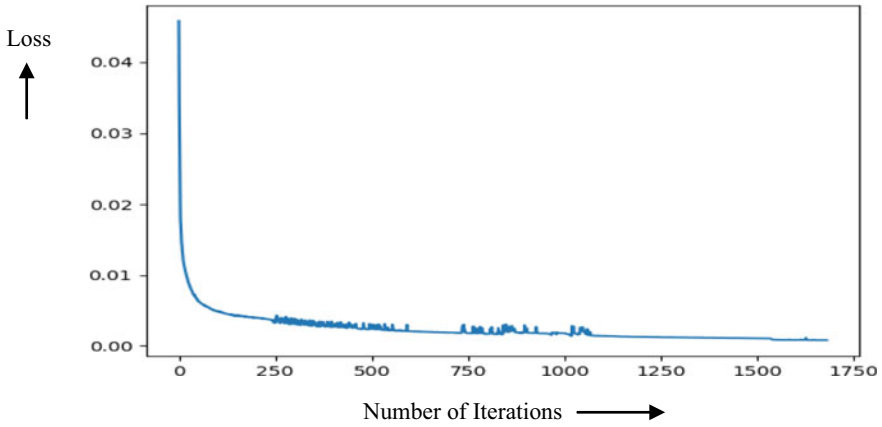


Fig. 4 Plot of loss with respect to number of iterations

If the user wants to retrain the network even after terminating the training of the network, then that option can also be chosen, i.e., retraining which will ask the user to give the name of the file where updated weights were already saved in the last training. After fetching the training data and updated weight parameters from the saved file, it will start retraining the network.

Generalization: After training of the network, it can be used for generalization where an unknown data or feature vector will be fed to the trained network. The desired class label will be produced using updated weight parameters. In generalization, new data will be fed which is not known to the trained network (Fig. 4).

5 Discussion and Conclusion

Literature consists of variants or different architectures of the same model. The paper presents the experiments performed on the different architecture with different variants like learning rate, number of hidden neurons number of epochs. Various observations were made while training MLP and CNN. Here, this paper tends to provide the insights of the results produced while evaluating our model on different architectures.

In the case of MLP, the design parameters like learning rate, weights, bias, and network topology are initialized. Weights and bias can be randomly initialized. Hit and trial method has been used to find the architecture with minimized cost. If the output of the architecture converges to the global minima, that architecture will be considered as the optimized architecture for the given problem. But if it stuck therein local minima thereupon the hit and trial method will be followed again by updating the design parameters of the network. Using Table 2, it has been interpreted that the

architecture with 0.1 learning rate and 100 nodes in the hidden layer gives minimized cost equals to 0.000676 with 89.78% testing accuracy, so this architecture can be considered as the optimized architecture in case of specific problem dataset.

The limitation of MLP is that these networks face overfitting problem oftentimes which over fits the network and misclassify the testing data but this problem can be solved by using regularization technique such as dropout. Sometimes, it also faces gradient vanishing problem which can be solved by using deep learning algorithms.

References

1. Sazli MH (2006) A brief review of feed-forward neural networks. Commun Fac Sci Univer Ankara 50(1):11–17
2. Sharma V, Rai S, Dev A (2012) A comprehensive study of artificial neural networks. Int J Adv Res Comput Sci Softw Eng 2(10)
3. Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance. Int J Remote Sens 28(5):823–870
4. Kamavisdar P, Saluja S, Agrawal S (2013) A survey on image classification approaches and techniques. Int J Adv Res Comput Commun Eng 2(1):1005–1009
5. Fagella D (2019) What is Machine Learning. https://emerj.com/ai-glossary-terms/what-is-machine-learning/. Last accessed 15 Apr 2019
6. Ramkumar T (2016) Analysis of multilayer perceptron machine learning approach in classifying protein secondary structures. Biomed Res 27:166–173
7. Ramchoun H, Amine M, Idrissi J, Ghanou Y, Ettaouil M (2016) Multilayer perceptron: architecture optimization and training. IJIMAI 4(1):26–30
8. Ertam F, Aydın G (2017) Data classification with deep learning using Tensorflow. In: 2017 international conference on computer science and engineering (UBMK). IEEE, pp 755–758
9. Campos PG, Oliveira EM, Ludermir TB, Araujo AF (2004) MLP networks for classification and prediction ith rule extraction mechanism. In: 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541), vol 2. IEEE, pp 1387–1392
10. Dua D, Graff C (2019) UCI machine learning repository. Irvine, CA, University of California, School of Information and Computer Science. http://archive.ics.uci.edu/ml

Brain Tumor Classification into High-Grade and Low-Grade Gliomas Using Adaboost

Sanjeet Pandey, Munesh Chandra Trivedi, Brijesh Bharadwaj, and Vineet Kr. Singh

Abstract Brain is recognized as one of the complex organs of the human body. Abnormal formation of cells may affect the normal functioning of the brain. These abnormal cells may belong to category of benign cells resulting in low-grade glioma or malignant cells resulting in high-grade glioma. The treatment plans vary according to grade of glioma detected. This results in need of precise glioma grading. As per World Health Organization, biopsy is considered to be gold standard in glioma grading. Biopsy is an invasive procedure which may contain sampling errors. Biopsy may also contain subjectivity errors. This motivated the clinician to look for other methods which may overcome the limitations of biopsy reports. Machine learning and deep learning approaches using MRI is considered to be the most promising alternative approach reported by scientist in literature. The presented work were based on the concept of AdaBoost approach which is an ensemble learning approach. The developed model was optimized w.r.t to two hyper parameters, i.e., number of estimators and learning rate keeping the base model fixed. The decision tree was used as a base model. The proposed developed model was trained and validated on BraTS 2018 dataset. The developed optimized model achieves reasonable accuracy in carrying out classification task, i.e., high-grade glioma vs. low-grade glioma.

Keywords High-grade glioma · Low-grade glioma · AdaBoost · Texture features · Feature selection

S. Pandey · V. Kr. Singh
MUIT, Lucknow, India

M. C. Trivedi (✉)
Department of CSE, NIT Agartala, IET, Agartala, India

B. Bharadwaj
Dr. RL Avadh University, Ayodhya, India

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2_41

1 Introduction

Brain is considered to be one of the complex organs of the body. If occurrence of uncontrolled division of cell takes place within the brain due to which abnormal formation of group of cells results in brain tumor. Tumor is considered to be life-threatening disease. This abnormal growth of the cell may affects the normal functioning (Fig. 1).

Brain tumors were majorly classified in low-grade tumor and high-grade tumor. Grade I and Grade II tumors are considered to be low-grade tumors, and Grade III and Grade IV tumors are considered to be high-grade tumor [2]. Low-grade tumor are considered to be non-cancerous or in other words, less aggressive in comparison to high-grade tumor. Exact causes of brain tumors are unknown till date, and researchers are conducting research to know the causes of brain tumor [3–7]. Some of the symptoms of brain tumor includes headache, difficulty in speaking, loss of movement etc. Interesting thing about brain tumor is that sometimes it does not shows the above-mentioned symptoms and can discovered accidently.

In order to detect the tumor, doctor may conduct investigations which may include imaging scans or biopsy or combination of both. Once tumor presence is confirmed, doctor may plan treatment and follow-required in process.

Magnetic resonance imaging is considered to be one of the favorite choice of investigation [2–5, 8–18]. Figure 2 shows the some of the conventional MRI sequences such as T2, FLAIR and T1 CE, respectively, with tumor.

Once tumor presence is confirmed in the MRI, clinician may plan biopsy to know the type and grade of the tumor. Sometimes repeated biopsies may be performed by the clinicians when tumor tissues were not enough to define the type or grade of tumor or if there was any confusions. Biopsy is an invasive procedure and may involve subjective and sampling errors. Errors in investigation procedure may affects the clinical treatment planning and follow-ups [2–5, 8–12].

Researchers were worked and still are working in the direction to address on questions like can invasive biopsies be replaced, can sampling errors may be reduced

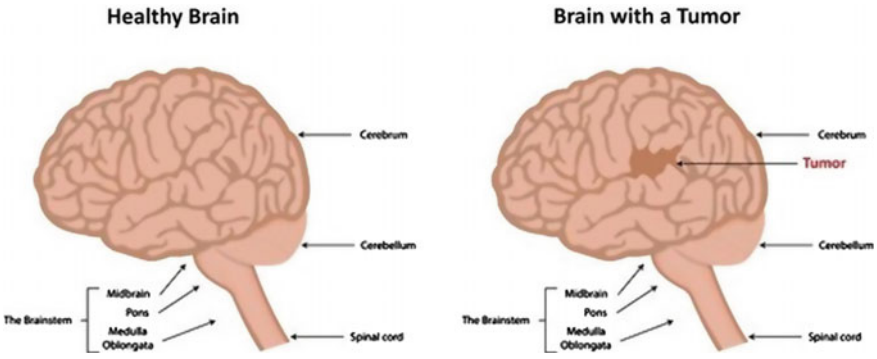


Fig. 1 Figure showing healthy brain and brain with tumor [1]

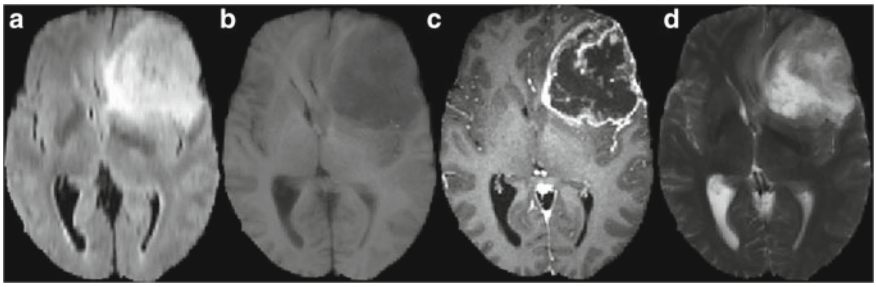


Fig. 2 Row-1 show MRI images (FLAIR). **a** T_1-weighted. **b** Post-contrast T1-weighted. **c** T_2-weighted. **d** Brain with tumor [8–10]

etc. Clinicians, scientist and engineers from cross-disciplinary areas are working in this direction. MRI investigations is considered to be non-invasive procedure [5, 11–13]. Quantitative features which were extracted from MRI were investigated as it is or with the help of machine leaning or deep leaning or transfer learning or any other procedure to identify the type and the grade of glioma. Positive results which were obtained with the help of machine learning or deep learning motivates the researcher to further investigate and improve the results in this direction. Some of the challenges which were mentioned by researchers in their findings were limited dataset, class imbalance error, subjectivity involve in tumor segmentation, cost and time.

In the proposed work, a hypothesis was presented which tries to differentiate the low-grade gliomas from high-grade glioma using the conventional MRI sequences using the texture features. AdaBoost algorithm was used to perform this classification task. Pearson correlation coefficient was used to select the features that will contribute in the classification task. Finally, a tenfold cross-validation was used to validate the trained model. Developed model is tested against the out of sample errors for recording the accuracy.

The rest of the paper was organized as follows: Sect. 2 discusses related work. Section 3 discusses the proposed classification approach. Section 4 discusses the obtained results on MATLAB 2020b platform, and Sect. 5 concludes the paper.

2 Related Work

This section describes the available literature work in the area of gliomas classification, i.e., HGG vs LGG. Faten et al. in their work used advanced sequences, i.e., diffusion tensor, perfusion etc., along with convention imaging in making differentiation between LGG and HGG [5]. Authors in their work [11] used the conventional MRI sequences along with advanced MRI sequences such as diffusion weighted imaging in classifying gliomas into LGG vs. HGG. The reported accuracy in their findings were 94.5%. Shoaib et al. in their work carried out the similar task and reported the accuracy equal to 80.65% [12]. A Vamvakas et al. in their studies reported

the classification accuracy equal to 95.5% [13]. They have used MRI conventional sequences, advanced sequences plus spectroscopy findings in their carried out study. Y. Yang et al. in their study used the concept of transfer learning in carrying out the classification task. They used MRI conventional sequences in their study. Their reported accuracy was 86.7% [14]. W. Chen et al. in their study investigated the role of radiomics in classification task, i.e., LGG vs. HGG [15]. Zurfi et al. [16] in their studies used 3D texture analysis in gliomas grading task with the help of machine learning. Authors [17, 18] in their work used the radiomics features which when fed as input to machine leaning algorithms to carry or gliomas grading task.

Although, several authors worked in this area and still research is going on. The major challenges mentioned by these authors in their manuscript were small dataset, reproducibility of results, globalized medical data, different acquisition protocols across different vendors, cost, subjectivity error etc. To address some of these issues, a globally publically available BraTS 2018 dataset has been used [8–10]. To reduce the cost in acquisition of advanced sequences, only conventional MRI sequences were used to carry out the desired classification task.

3 Proposed Work

This section explains the proposed work. BraTS 2018 dataset was used for carrying out the classification task. The dataset contains 210 high-grade glioma cases and 75 low-grade glioma cases. For every case, the dataset contains T_1, post-contrast T_1, T_2, fluid attenuated inversion recovery (FLAIR) sequences. The dataset belongs to 19 different centers. The dataset was annotated into four labels:

Label-0 otherwise.
Label-1 Non-enhancing tumor and necrotic region.
Label-2 Edema.
Label-4 Enhancing tumor.

Every sequence in BraTS 2018 dataset co-registered and interpolated. Texture features were extracted from region of interest (ROI) with the help of pyradiomics using python [6, 7]. Label 1 and label 4 were combined to form ROI. A total 104 features were computed which belongs to seven different classes, i.e., shape-based (2D), gray-level matrix (co-occurrence, run length, dependence and size zone matrix) and neighboring gray tone difference matrix. Feature selection were made with the help of Pearson correlation coefficient method. Features were normalized using the concept of z-score. Finally, 49 features were selected out of 104 computed features.

AdaBoost algorithm was used for carrying out the classification task. AdaBoost algorithm combines various weak learners to form a strong learner based on ensemble concept [19]. A tenfold cross validation were performed to finally validate the model. The mean accuracy was calculated across the tenfolds by developing different models by varying the number of estimators and learning rate (0.001–1). Decision tree was used as base estimator.

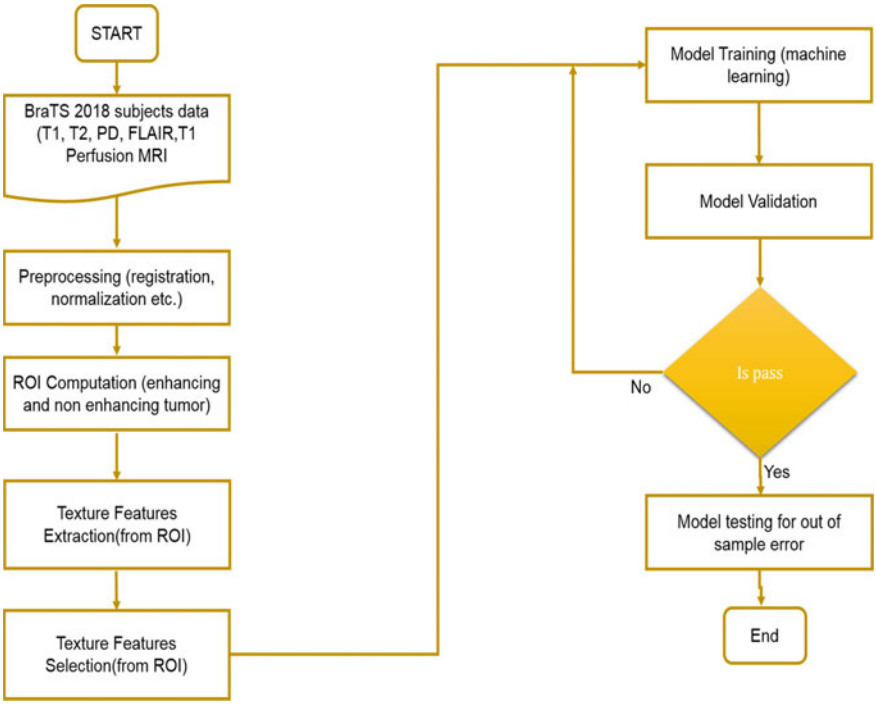


Fig. 3 Figures shows the systematic diagram of proposed concept (classification model development)

4 Results

Learning rate was varied from 0.001 to 1 and number of estimators were varied from 10 to 400. It was noted further increasing in number of estimates shows no improvement in accuracy and hence not shown in Fig. 4. From Fig. 4, it was observed that model performs better when number of estimators were equal to 150. From Fig. 5, it was observed that model performs well when learning rate was equal to 0.1. Final model was developed keeping the hyper parameters, i.e., learning rate equals to 0.1 and number of estimators equals to 150. Developed model achieves the accuracy equals to 86.3% in classifying the tumor into high grade vs. low grade (Fig. 3; Table 1).

5 Conclusion

The whole paper was organized in four major sections: introduction, related work, proposed work and simulation results. Introduction section briefly explains the need

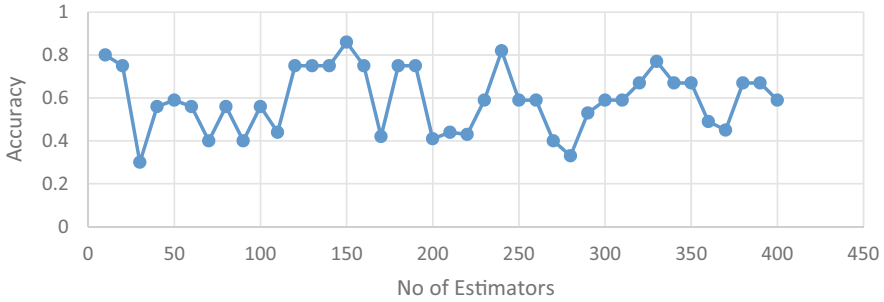


Fig. 4 Shows the accuracy achieved by models by varying the number of estimators keeping the base model fixed

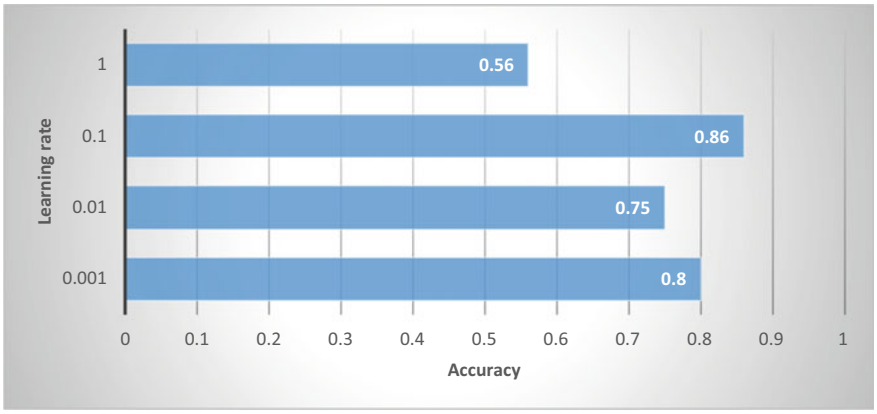


Fig. 5 Shows the accuracy achieved by models by varying the learning rate keeping the base model fixed

Table 1 Showing the optimized classifier performance

Base classifier	Number of estimator	Learning rate	Accuracy	Sensitivity	Specificity
Decision tree	150	0.1	86.3	84.2	90.1

of brain tumor diagnosis. This section also explains the limitation of biopsy procedure and hence establishes need of precise glioma classification. Sect. 2, i.e., related work discusses some of the recent work carried out by clinician and scientist in the area of glioma classification. In Sect. 3, proposed work has been discussed. AdaBoost was used as an underlying concept to develop the model to carry out the designated task. Hyperparameters were optimized and cross validated. Finally, model was developed over these optimized hyperparameters keeping the base estimator same. In result section, only the final optimized model accuracy were reported along with sensitivity and specificity. The results show the model achieved the reasonable

accuracy in classifying high-grade glioma from low-grade glioma. Hence, concludes the presented paper.

Acknowledgements We would like to thanks our professor for their kind guidance and financial support in order to carry out this task. We also thank our parents and friends for their moral support and encouragement. We also thanks almighty god without his blessings this work is not possible.

References

1. Brain Tumours V Brain Aneurysms: Understanding the Difference | Medanta https://www.medanta.org/patient-education-blog/brain-tumours-v-brain-aneurysms-understanding-the-difference/. Accessed 22 Mar 2021
2. "Gliomas Department of Neurology." [Online]. Available https://www.columbianeurology.org/neurology/staywell/document.php?id=42006. Accessed 25 Jul 2020
3. All About Adult Gliomas | OncoLink." [Online]. Available https://www.oncolink.org/cancers/brain-tumors/all-about-adult-gliomas. Accessed 25 Jul 2020
4. "Gliomas | Johns Hopkins Medicine." [Online]. Available https://www.hopkinsmedicine.org/health/conditions-and-diseases/gliomas. Accessed 25 Jul 2020
5. Fawzy FM, Almassry HN, Ismail AM (2016) Preoperative glioma grading by MR diffusion and MR spectroscopic imaging. Egypt J Radiol Nucl Med. 47(4):1539–1548
6. Vallières M et al (2015) A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 60(14):5471–5496. https://doi.org/10.1088/0031-9155/60/14/5471
7. Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillon-Robin JC, Pieper S, Aerts HJWL (2017) Computational radiomics system to decode the radiographic phenotype. Cancer Res 77(21):e104–e107. https://doi.org/10.1158/0008-5472.CAN-17-0339 <https://doi.org/10.1158/0008-5472.CAN-17-0339>
8. Menze BH, Jakab A et al (2015) The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 34(10):1993–2024
9. Bakas S, Akbari H et al (2017) Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Nat Sci Data. 4:170117
10. Bakas S, Akbari H, et al (2018) Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. arXiv preprint arXiv:1811.02629
11. Zhang X, et al (2017) Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features
12. Shoaib Y et al (2019) Role of diffusion and perfusion magnetic resonance imaging in predicting the histopathological grade of gliomas—a prospective study. Asian J Neurosurg 14(1):47
13. Vamvakas et al (2019) Imaging biomarker analysis of advanced multiparametric MRI for glioma grading. Phys Medica
14. Yang Y et al (2018) Glioma grading on conventional MR images: a deep learning study with transfer learning. Front Neurosci 12:804
15. Chen W, Liu B, Peng S, Sun J, Qiao X (2018) Aided grading of gliomas combining automatic segmentation and radiomics
16. Al-Zurfi N, Meziane F, Aspin R (2019) A computer-aided diagnosis system for glioma grading using three dimensional texture analysis and machine learning in MRI brain tumour. In: 2019 3rd international conference on bio-engineering for smart technologies (BioSMART), pp 1–5
17. Zhan T, Feng P, Hong X, Lu Z, Xiao L, Zhang Y (2017) An automatic glioma grading method based on multi-feature extraction and fusion. Technol Health Care 25(S1):S377–S385

18. “Glioma Classification of MR Brain Tumor Employing Machine Learning. [Online]. Available: https://www.researchgate.net/publication/335340018_Glioma_Classification_of_MR_Brain_Tumor_Employing_Machine_Learning. Accessed 26 Jul 2020
19. Boosting and AdaBoost for Machine Learning. https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/. Accessed 21 Mar 2021

Author Index

X.-Z. Gao et al. (eds.), *Advances in Computational Intelligence and Communication Technology*, Lecture Notes in Networks and Systems 399,
https://doi.org/10.1007/978-981-16-9756-2